

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 23, 2003, 13:59:02 ; Search time 53.406 Seconds
(without alignments)
825.861 Million cell updates/sec

Title: US-10-056-052A-4
Perfect score: 1732
Sequence: 1 MVAADAPAAAGTDITNQLTNV.....NGSGSGDGIDKPVVPEQPDE 331

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues
Number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database :			
A_Geneseq_101002:*			
1:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*		
2:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*		
3:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*		
4:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*		
5:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*		
6:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*		
7:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*		
8:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*		
9:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*		
10:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*		
11:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*		
12:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*		
13:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*		
14:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*		
15:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*		
16:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*		
17:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*		
18:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*		
19:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*		
20:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*		
21:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*		
22:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*		
23:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*		

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES			
Result No.	Score	Query Match Length DB ID	Description
1	1732	100.0 345 19	AAW31555 Fibronectin-bindin
2	1727	99.7 933 21	AAV58435 Staphylococcus aur
3	1727	99.7 933 22	AAW69508 Staphylococcus aur
4	1723	99.5 936 18	AAW89801 Staphylococcus aur
5	1518	87.6 1021 22	AAU33975 Staphylococcus aur
6	1518	87.6 1021 22	AAU36951 Staphylococcus aur
7	1433	82.7 496 23	AAU75490 S. aureus antigeni
8	409	23.6 767 22	AAU34403 Staphylococcus aur
9	409	23.6 767 22	AAU37547 Staphylococcus aur
10	409	23.6 940 11	AAU07070 Fibronectin-bindin

PA (QUEE-) QUEEN ELIZABETH COLLEGE DUBLIN.
XX
PI Foster TJ, McDevitt DL;
XX
XX WPI; 2000-096389/08.
DR N-PSDB; AA255832.
XX
XX Nucleic acid encoding the fibrinogen-binding protein of Staphylococcus
PT aureus, useful for treatment or prevention of infections .
XX
PS Disclosure; Fig 2A-1-4; 35pp; English.
XX
XX This sequence represents the Staphylococcus aureus
CC fibrinogen-binding protein, ClfA. ClfA is an important receptor
CC involved in S. aureus colonisation of indwelling medical devices (e.g.,
CC catheters, artificial heart valves). Shortly after implantation, the
CC surfaces of medical devices become coated with host plasma and matrix
CC proteins such as fibrinogen and fibronectin, and there is considerable
CC evidence to suggest that bacterial adherence to fibrinogen/fibrin is
CC important in the initiation of device-related infection. The
CC fibrinogen-binding region of ClfA is thought to reside between residues
CC 332 and 550 in a region designated A. The protein also contains a
CC repeated region (region R) comprising 154 repeats of the dipeptide
CC Ser-Asp, and the C-terminus contains features present in surface
CC proteins of other Gram positive bacteria that are responsible for
CC anchoring the protein to the cell wall and cell membrane. ClfA, or its
CC fragments, may be used to block S. aureus colonisation of wounds, to
CC prevent adherence of S. aureus to indwelling medical devices, as
CC vaccines to protect against S. aureus infection (e.g., mastitis in
CC ruminants), to raise specific antibodies, and for diagnosis (by
CC agglutination or immunoassay). The specific antibodies
CC are used for passive immunisation, to block infection of wounds or
CC adhesion of S. aureus and for diagnosis. Nucleotides encoding ClfA and
CC its fragments may be used as diagnostic probes.
XX
SQ Sequence 933 AA;

Query Match 99.7%; Score 1727; DB 21; Length 933;
Best Local Similarity 100.0%; Pred. No. 9.5e-116;
Matches 330; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAADAPAAAGTDITNQLTNVTVGIDSGTTVYPHQAGYVKLVNKGFSVPNSAVKGDTEFKITVP 61
Db |
221 VAADAPAAAGTDITNQLTNVTVGIDSGTTVYPHQAGYVKLVNKGFSVPNSAVKGDTEFKITVP 280

QY 62 KELNLNGVTSTAKVPPIMAGDQVLANGVIDSDGNVIYFTFDYVNTKDDVKATLTMPAYID 121
|
281 KELNLNGVTSTAKVPPIMAGDQVLANGVIDSDGNVIYFTFDYVNTKDDVKATLTMPAYID 340

122 PENVKKTGNVTLATGIGSTTANKTVLVVDYEKYGKFYNLSIKGTIDQIDKTNNTYRQTIYV 181
|
341 PENVKKTGNVTLATGIGSTTANKTVLVVDYEKYGKFYNLSIKGTIDQIDKTNNTYRQTIYV 400

Db |
182 NPSGDNVIAPVLTGNLKPNTDSNALIDQQNTSIKVKYKVDNAADLSESYFVNPFEDVTN 241
|
401 NPSGDNVIAPVLTGNLKPNTDSNALIDQQNTSIKVKYKVDNAADLSESYFVNPFEDVTN 460

QY |
242 SVNITFPNPQYKVEFNTPDQITTPYIVVNVNGHIDPNKSGDLALRSTLYGYNNSNIWRS 301
|
461 SVNITFPNPQYKVEFNTPDQITTPYIVVNVNGHIDPNKSGDLALRSTLYGYNNSNIWRS 520

QY 302 MSWDNEVAFNNGSGSGDGIDKPVVPEQPDE 331
|
521 MSWDNEVAFNNGSGSGDGIDKPVVPEQPDE 550

Db |

RESULT 3
AAB69508
ID AAB69508 standard; Protein; 933 AA.
XX
AC AAB69508;
XX
DT 23-APR-2001 (first entry)

XX Staphylococcus aureus ClfA protein.
DE
XX
XX Staphylococcus aureus; clfa; antibiotic; vaccine;
KW fibrinogen binding protein; bacterial infection; mastitis.
XX
OS Staphylococcus aureus.
XX
PN US6177084-B1.
XX
PD 23-JAN-2001.
XX
PF 19-OCT-1999; 99US-0421868.
XX
PR 22-AUG-1994; 94US-0293728.
XX
PA (QUEE-) QUEEN ELIZABETH COLLEGE DUBLIN.
XX
PI Foster TJ, McDevitt DL;
XX
XX WPI; 2001-181608/18.
DR N-PSDB; AAF58593.
XX
PT Novel Staphylococcus aureus fibrinogen binding protein useful as
PT vaccine to prevent infection, promote wound healing, block adherence to
PT indwelling medical devices and for diagnosing staphylococcus aureus
PT infection .
XX
PS Claim 5; Fig 2; 30pp; English.
XX
CC The present sequence is a novel Staphylococcus aureus fibrinogen
CC binding protein. It is useful as a vaccine to protect against human
CC and animal infections caused by S.aureus, such as against mastitis, to
CC block S.aureus from colonising and infecting a wound, to block adherence
CC of S.aureus to indwelling medical devices such as catheters, replacement
CC heart valves and cardiac assist devices. The protein can be used to
CC diagnose bacterial infections. Polyclonal and monoclonal antibodies
CC raised against the fibrinogen binding protein are useful for passively
CC immunising against infections caused by S.aureus, to prevent infection
CC of a wound and to diagnose bacterial infections.
XX
SQ Sequence 933 AA;

Query Match 99.7%; Score 1727; DB 22; Length 933;
Best Local Similarity 100.0%; Pred. No. 9.5e-116;
Matches 330; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAADAPAAAGTDITNQLTNVTVGIDSGTTVYPHQAGYVKLVNKGFSVPNSAVKGDTEFKITVP 61
|
221 VAADAPAAAGTDITNQLTNVTVGIDSGTTVYPHQAGYVKLVNKGFSVPNSAVKGDTEFKITVP 280

QY 62 KELNLNGVTSTAKVPPIMAGDQVLANGVIDSDGNVIYFTFDYVNTKDDVKATLTMPAYID 121
|
281 KELNLNGVTSTAKVPPIMAGDQVLANGVIDSDGNVIYFTFDYVNTKDDVKATLTMPAYID 340

QY 122 PENVKKTGNVTLATGIGSTTANKTVLVVDYEKYGKFYNLSIKGTIDQIDKTNNTYRQTIYV 181
|
341 PENVKKTGNVTLATGIGSTTANKTVLVVDYEKYGKFYNLSIKGTIDQIDKTNNTYRQTIYV 400

QY 182 NPSGDNVIAPVLTGNLKPNTDSNALIDQQNTSIKVKYKVDNAADLSESYFVNPFEDVTN 241
|
401 NPSGDNVIAPVLTGNLKPNTDSNALIDQQNTSIKVKYKVDNAADLSESYFVNPFEDVTN 460

QY 242 SVNITFPNPQYKVEFNTPDQITTPYIVVNVNGHIDPNKSGDLALRSTLYGYNNSNIWRS 301
|
461 SVNITFPNPQYKVEFNTPDQITTPYIVVNVNGHIDPNKSGDLALRSTLYGYNNSNIWRS 520

QY 302 MSWDNEVAFNNGSGSGDGIDKPVVPEQPDE 331
|
521 MSWDNEVAFNNGSGSGDGIDKPVVPEQPDE 550

Db |

RESULT 4

AAW89801	ID	AAW89801 standard; Protein; 936 AA.
XX	XX	
AC	AAW89801;	
XX	16-MAR-1999	(first entry)
DT	XX	
XX	DE	Staphylococcus aureus protein SEQ ID #5249.
XX	XX	
KW	Computer readable medium; vaccine; S.aureus infection; immunodetection;	
KW	cellulitis; eyelid infection; food poisoning; osteomyelitis; therapy;	
KW	skin infection; surgical wound infection; scalded skin syndrome;	
XX	toxic shock syndrome.	
.XX		
OS	Staphylococcus aureus.	
XX	XX	
PN	EP786519-A2.	
XX	30-JUL-1997.	
PD	XX	
XX	07-JAN-1997;	97EP-0100117.
PR	05-JAN-1996;	96US-0009861.
XX	(HUMA-) HUMAN GENOME SCI INC.	
PA	Barash SC, Choi GH, Dillon PJ, Fannon MR, Kunsch CA;	
PI	Rosen CA;	
PI	XX	
XX	WPI; 1997-374922/35.	
DR	XX	
XX	Polynucleotide(s) and proteins derived from Staphylococcus aureus	
PT	stored on computer readable medium and used in the production of	
PT	anti-S.aureus vaccines	
PT	XX	
XX	Claim 23; Page 3255-3258; 3271pp; English.	
PS	XX	
XX	This sequence represents a Staphylococcus aureus protein sequence of the	
CC	invention. The DNA sequences encoding the S.aureus proteins are recorded	
CC	on a computer readable medium, preferably selected from a floppy or hard	
CC	disk, random access memory (RAM), read-only memory (ROM) or CD-ROM.	
CC	Homology searches using the S.aureus DNA sequences allows putative	
CC	functions to be assigned so that protein-encoding or regulatory regions	
CC	of commercial, therapeutic or industrial importance can be obtained.	
CC	Specifically, sequences which are likely to encode antigens have been	
CC	identified and these polypeptides can be used in a vaccine composition	
CC	against S.aureus infection. The polypeptides can also be used in a kit	
CC	for the immunodetection of S.aureus in a sample. S.aureus is implicated	
CC	in numerous human diseases, including cellulitis, eyelid infections, food	
CC	poisoning, osteomyelitis, skin and surgical wound infections, scalded	
CC	skin syndrome, toxic shock syndrome, etc. Organisms transformed with the	
CC	DNA sequences can be used for recombinant production of the polypeptides.	
CC	The new DNA sequences (and their fragments) are useful as primers or	
CC	probes for isolating homologues of any of the 5191 S.aureus DNA sequences	
CC	contained on the computer readable medium.	
XX		
SQ	Sequence	936 AA;
	Query Match	99.5%; Score 1723; DB 18; Length 936;
	Best Local Similarity	99.7%; Pred. No. 1.9e-115;
	Matches 329; Conservative	0; Mismatches 1; Indels 0; Gaps 0
QY	2	VAADAPAGTDITNQLTNVTGIDSGTTVYPHQAGYVKLVYGFSPNSAVKGDTFKITVP 61
Db	230	VAADAPVAGTDITNQLTNVTGIDSGTTVYPHQAGYVKLVYGFSPNSAVKGDTFKITVP 289
QY	62	KELNLNGVTSTAKVPPIMAGDQVLANGVIDSDGNVIYFTFDYVNTKDDVKATLTMPAYID 121
Db	290	KELNLNGVTSTAKVPPIMAGDQVLANGVIDSDGNVIYFTFDYVNTKDDVKATLTMPAYID 349
QY	122	PENVKKTGNVTLATGIGSTTANKTVLDYKEYGKFYNLSIKGTIDQIDKTNNTYRQTIYV 181
Db	350	PENVKKTGNVTLATGIGSTTANKTVLDYKEYGKFYNLSIKGTIDQIDKTNNTYRQTIYV 409

QY	182	NPSGDNVIAPVLGTNLKPNTDSNALIDQQNTSIKVYKVDNAADLSesyFVNPFEDVTN	241
Db	410	NPSGDNVIAPVLGTNLKPNTDSNALIDQQNTSIKVYKVDNAADLSesyFVNPFEDVTN	469
QY	242	SVNITFPNPQYKVEFNTPDDQITTPYIVVWNGHIDPNSKGDALRSTLYGYNsNIWRS	301
Db	470	SVNITFPNPQYKVEFNTPDDQITTPYIVVWNGHIDPNSKGDALRSTLYGYNsNIWRS	529
QY	302	MSWDNEVAfNNGSGGDGIDKPVVPDE	331
Db	530	MSWDNEVAfNNGSGGDGIDKPVVPDE	559
 RESULT 5 AAU33975			
ID	AAU33975	standard; Protein; 1021 AA.	
XX	AC	AAU33975;	
XX	DT	14-FEB-2002 (first entry)	
XX	DE	Staphylococcus aureus cellular proliferation protein #251.	
XX	DE	Antisense; prokaryotic cellular proliferation protein;	
KW	KW	antibiotic; antibacterial; drug design.	
OS	OS	Staphylococcus aureus.	
XX	PN	WO200170955-A2.	
XX	PD	27-SEP-2001.	
XX	PF	21-MAR-2001; 2001WO-US09180.	
XX	PR	21-MAR-2000; 2000US-191078P.	
XX	PR	23-MAY-2000; 2000US-206848P.	
XX	PR	26-MAY-2000; 2000US-207727P.	
XX	PR	23-OCT-2000; 2000US-242578P.	
XX	PR	27-NOV-2000; 2000US-253625P.	
XX	PR	22-DEC-2000; 2000US-257931P.	
XX	PR	16-FEB-2001; 2001US-269308P.	
PA	PA	(ELIT-) ELITRA PHARM INC.	
PI	PI	Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;	
PI	PI	Yamamoto RT, Xu HH;	
XX	DR	WPI; 2001-611495/70.	
DR	DR	N-PSDB; AAS51834.	
XX	PT	New polynucleotides for the identification and development of	
PT	PT	antibiotics, comprise sequences of antisense nucleic acids -	
XX	PS	Example 3; Seq ID NO 5471; 511pp; English.	
XX	CC	The invention relates to antisense inhibitors of genes essential to	
CC	CC	prokaryotic cellular proliferation, their use in identifying the	
CC	CC	genes, their use in the discovery of novel antibiotics, the essential	
CC	CC	genes themselves and the encoded proteins. The prokaryotes used are	
CC	CC	Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella	
CC	CC	pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The	
CC	CC	invention is also useful for the identification of potential new targets	
CC	CC	for antibiotic development. The antisense nucleic acids can also be used	
CC	CC	to identify proteins used in proliferation, to express these proteins,	
CC	CC	and to obtain antibodies capable of binding to the expressed proteins.	
CC	CC	The proteins can be used to screen compounds in rational drug discovery	
CC	CC	programmes. The antisense nucleic acid sequence is also useful to screen	
CC	CC	for homologous nucleic acids which are required for cell proliferation in	
CC	CC	a wide variety of organisms. The present sequence represents an	
CC	CC	essential prokaryotic cellular proliferation protein.	
CC	CC	Note: The sequence data for this patent did not form part	
CC	CC	of the printed specification, but was obtained in electronic	

Db 155 VVEETKATGTDVTK-----VEVEEGSEIVGHKQDTNVVNPHPNAERVTLKYKWKFGEGIK 209

Qy 52 KGDTFKITVPKELNLNGVTSTAKVPPIMAGD-QVLANGVIDSDGNVIYFTDYVNTKDDV 110

Db 210 AGDYFDFTLSDNVETHGISTLRKKVPEIKSTOGQVMTGEIIGERKVRVYTFKEYVQEKKDL 269

Qy 111 KATLTMPAYIDPENVKKTGNVTLATGIGSTTANKTVLVVDY-----EKYGFYNLSIKGTI 165

Db 270 TAELSLNLFIIDPTTQKGNQNVKVEKLGGETTVSKIFNIQYLGGVDRNWG----VTANGRI 325

Qy 166 DQIDKTNNTYRQTIYVNPSPGDNVIAPVLGTNL----KPNVDSNALIDQQNTSIKVKVDN 221

Db 326 DTLNKVDGKFSHFAYMKPNQNSLSSVTVTGQVTKGNKPGVN-----NPTVKVYKHIG 377

Qy 222 AADLSESYFVNPNEN---FEDVTVNSVNITFPNPNOYKVEFNTPDQITTPYIVVVNGHIDP 278

Db 378 SDDLAEVYAKLDDVSKFEDVTDNMSLDFDTNGGYSLNFNNLDQ--SKNYVIKYEYGYDS 435

Qy 279 NSKGDALALRSTLYGYNYSNIWRSMWDNEVAFNNGSGSGDIDK---PVV 325

436 NA-SNLEFQTHLFGYNYNYTSLNLTWKNGVAFYSNNAQDGDKLKEPII 484

RESULT 9

AAU37547

ID AAU37547 standard; Protein; 767 AA.

XX AAU37547;

DT 14-FEB-2002 (first entry)

DE Staphylococcus aureus cellular proliferation protein #1717.

KW Antisense; prokaryotic cellular proliferation protein;
antibiotic; antibacterial; drug design.

OS Staphylococcus aureus.

PN WO200170955-A2.

PD 27-SEP-2001.

PF 21-MAR-2001; 2001WO-US09180.

PR 21-MAR-2000; 2000US-191078P.

PR 23-MAY-2000; 2000US-206848P.

PR 26-MAY-2000; 2000US-207727P.

PR 23-OCT-2000; 2000US-242578P.

PR 27-NOV-2000; 2000US-253625P.

PR 22-DEC-2000; 2000US-257931P.

PR 16-FEB-2001; 2001US-269308P.

PA (ELIT-) ELITRA PHARM INC.

PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
Yamamoto RT, Xu HH;

XX WPI; 2001-611495/70.

DR N-PSDB; AAS55406.

PT New polynucleotides for the identification and development of
antibiotics, comprise sequences of antisense nucleic acids -

XX Example 3; Seq ID No 13140; 511pp; English.

CC The invention relates to antisense inhibitors of genes essential to
prokaryotic cellular proliferation, their use in identifying the
genes, their use in the discovery of novel antibiotics, the essential
genes themselves and the encoded proteins. The prokaryotes used are
Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
invention is also useful for the identification of potential new targets
for antibiotic development. The antisense nucleic acids can also be used

CC to identify proteins used in proliferation, to express these proteins,
and to obtain antibodies capable of binding to the expressed proteins.
CC The programmes. The antisense nucleic acid sequence is also useful to screen
for homologous nucleic acids which are required for cell proliferation in
a wide variety of organisms. The present sequence represents an
essential prokaryotic cellular proliferation protein.
CC Note: The sequence data for this patent did not form part
of the printed specification, but was obtained in electronic
format directly from WIPO at
ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 767 AA;

Qy Query Match 23.6%; Score 409; DB 22; Length 767;

Best Local Similarity 28.9%; Pred. No. 4.2e-21;

Matches 101; Conservative 72; Mismatches 131; Indels 46; Gaps 11;

Qy 2 VAADAPAAAGTDITNQLTNVTVGIDSGT-----TVYPHQAGYVKLNYGFSVPNSAV 51

Db 155 VVEETKATGTDVTK-----VEVEEGSEIVGHKQDTNVVNPHPNAERVTLKYKWKFGEGIK 209

Qy 52 KGDTFKITVPKELNLNGVTSTAKVPPIMAGD-QVLANGVIDSDGNVIYFTDYVNTKDDV 110

Db 210 AGDYFDFTLSDNVETHGISTLRKKVPEIKSTOGQVMTGEIIGERKVRVYTFKEYVQEKKDL 269

Qy 111 KATLTMPAYIDPENVKKTGNVTLATGIGSTTANKTVLVVDY-----EKYGFYNLSIKGTI 165

Db 270 TAELSLNLFIIDPTTQKGNQNVKVEKLGGETTVSKIFNIQYLGGVDRNWG----VTANGRI 325

Qy 166 DQIDKTNNTYRQTIYVNPSPGDNVIAPVLGTNL----KPNVDSNALIDQQNTSIKVKVDN 221

Db 326 DTLNKVDGKFSHFAYMKPNQNSLSSVTVTGQVTKGNKPGVN-----NPTVKVYKHIG 377

Qy 222 AADLSESYFVNPNEN---FEDVTVNSVNITFPNPNOYKVEFNTPDQITTPYIVVVNGHIDP 278

Db 378 SDDLAEVYAKLDDVSKFEDVTDNMSLDFDTNGGYSLNFNNLDQ--SKNYVIKYEYGYDS 435

Qy 279 NSKGDALALRSTLYGYNYSNIWRSMWDNEVAFNNGSGSGDIDK---PVV 325

Db 436 NA-SNLEFQTHLFGYNYNYTSLNLTWKNGVAFYSNNAQDGDKLKEPII 484

RESULT 10

AAR07070

ID AAR07070 standard; protein; 940 AA.

XX AAR07070;

DT 21-FEB-1991 (first entry)

DE Fibronectin-binding protein encoded by gene 2, fnbB.

KW Wound healing; immunisation; mastitis; fnbA; ss.

OS Staphylococcus aureus.

XX Key Location/Qualifiers

FT Peptide 1..36

FT /label=signal peptide

XX EP397633-A.

XX 14-NOV-1990.

PF 04-MAY-1990; 90EP-0850166.

PR 14-NOV-1990; 90SE-0001687.

XX (ALFA) ALFA-LAVAL AGRI INT.

PI Hook M, Jonsson KL, Lindberg KM, Signas LC;

XX

XX FH Key Location/Qualifiers
FT Misc-difference 14 /note= "Position encoded by TAG stop codon"
FT Misc-difference 33 /note= "Position encoded by TGA stop codon"
FT Misc-difference 964 /note= "Position encoded by TAA stop codon"
FT Misc-difference 980 /note= "Position encoded by TAG stop codon"
FT Misc-difference 989 /note= "Position encoded by TAA stop codon"
XX PN WO200012689-A1.
XX PD 09-MAR-2000.
XX PF 31-AUG-1999; 99WO-US19728.
XX 31-AUG-1998; 98US-0098443.
XX 25-JAN-1999; 99US-0117119.
XX (QUEE-) QUEEN ELIZABETH COLLEGE DUBLIN.
PA (TEXA) UNIV TEXAS A & M SYSTEM.
XX Foster TJ, Hook M, Davis S, Hartford O, McCrear K, Ni Eidhin D;
WPI; 2000-256637/22.
DR N-PSDB; AA293534.
XX Recombinant or synthetic proteins from coagulase-negative staphylococci
PT useful for prevention, treatment and diagnosis of staphylococcal
PT infections bind soluble and immobilized fibrinogen
XX Claim 8; Figure 3; 104pp; English.
XX Isolated Staphylococcus Sdr cell wall proteins which bind both
CC soluble and immobilized fibrinogen are useful for treating or
CC preventing coagulase-negative staphylococcal infection such as
CC septicemia, osteomyelitis or endocarditis, and for inducing immune
CC responses in patients. The cell wall proteins are also useful for
CC reducing coagulase-negative staphylococci infection of indwelling
CC medical devices such as vascular grafts, vascular stents,
CC intravenous catheters, artificial heart valves and cardiac assist
CC devices. The cell wall associated proteins are able to inhibit
CC staphylococcal adhesion to immobilised extracellular matrix or host
CC cells present on the surface of implanted biomaterials.
Sequence 991 AA;
Query Match 20.4%; Score 353; DB 21; Length 991;
Best Local Similarity 30.6%; Pred. No. 6.1e-17;
Matches 101; Conservative 60; Mismatches 139; Indels 30; Gaps 14;
QY 13 ITNQLTNVTVGI-DSGTTVYPHQAGYVVKLNYGFSVPNSAVKGDTPKITYPKELNLNGVTS 71
Db 317 VTDO--SITEGYDDSDGIKAHDAENLIYDVTFEVDKVKSGDTMTVNIDKNTVPSDLTD 374
QY 72 TAKVPPIMAGD-QVLANGVID-SDGNVIYTFDYVNTKDDVKATLTMPAYIDPENVKKTG 129
Db 375 SFAIPKIKDNGSEIIATGTYDNTNKQITYTFTDYVDKYENIKAKHLKLTSDYIDKSKVPNNN 434
QY 130 ---NVTLATGIGSTTANKTVLVDYKYGKFNLSIKGTIDQIDKTNNTYRQTIYVNPSCD 186
Db 435 TKLDVEYKTALESS--VNKTITVEYQKPNENRTANLQSMFTNIDTKNHTVEQTIYINPL-- 490
QY 187 NVIAPVLGNLKPNTDSNALIDQNTSIKVKYVDNAADLSESYFV-NPENFEDVTNSVNI 245
Db 491 RYSAKETNVNISNGDEGSTIIDDSTIIKVKYVGDNQNLPSNRIYDYSEYEDVTNDYYA 550
QY 246 TFPNPQYKVEFNTPDQITTPYIVVNGHIDPNKSGD-----LALRSTLYGYNISII 298
Db 551 QLGNNNDVNI NFG----NIDSPYIIKVISKYDPN-KDDYTTIQQVTMTQTTINEYTG-- 603

QY 299 WRMSWDNEVAFNNGSGDGDIDKPVVPEQ 328
Db 604 FRTASYDNTIAFSTSSGGQGG-DLP--PEK 630
RESULT 13
AAY70120
ID AAY70120 standard; Protein; 991 AA.
XX AAY70120;
AC AAY70120;
XX 06-JUN-2000 (first entry)
XX Staph. epidermidis serine-aspartate repeat region protein SdrG.
DE KW Multicomponent vaccine; immunostimulatory; antibacterial; MSCRAMM;
XX microbial surface components recognising adhesive matrix molecules;
KW collagen binding protein; CBP; CNA; fibrinogen binding protein;
KW Clumping factor A; ClfA; Clumping factor B; ClfB; FnBP;
KW fibronectin binding protein; Staphylococcus infection;
KW serine-aspartate repeat region protein; SDR protein; SdrG.
XX Staphylococcus epidermidis.
OS Key Location/Qualifiers
XX Misc-difference 14 /note= "Encoded by in-frame stop codon TAG".
XX Misc-difference 33 /note= "Encoded by in-frame stop codon TGA".
XX Misc-difference 964 /note= "Encoded by in-frame stop codon TAA".
XX Misc-difference 980 /note= "Encoded by in-frame stop codon TAG".
XX Misc-difference 989 /note= "Encoded by in-frame stop codon TAA".
XX WO200012131-A1.
XX 09-MAR-2000.
XX 31-AUG-1999; 99WO-US19727.
XX 31-AUG-1998; 98US-0098439.
XX (INHI-) INHIBITEX INC.
PA (TEXA) UNIV TEXAS A & M SYSTEM.
PA (QUEE-) QUEEN ELIZABETH COLLEGE DUBLIN.
XX Patti JM, Foster TJ, Hook M;
PI WPI; 2000-237781/20.
XX N-PSDB; AA251202.
XX Composition used for generating immune response or for inhibiting
PT microbial colonization in an animal comprises antibodies that bind
PT collagen binding protein, fibrinogen binding protein and, optionally,
PT fibronectin binding protein
XX Claim 8; Fig 4; 115pp; English.
XX The patent discloses multicomponent vaccines containing selected
CC combinations of bacterial binding proteins termed MSCRAMM (microbial
CC surface components recognising adhesive matrix molecules) or their
CC antibodies. A vaccine composition is provided that includes collagen
CC binding protein or peptide, e.g. CNA, a fibrinogen binding protein
CC preferably Clumping factor A (ClfA) or Clumping factor B (ClfB),
CC and optionally a fibronectin binding protein e.g. FnBP-A.
CC The vaccines are useful for imparting protection against a broad
CC spectrum of Staphylococcal strains and for inhibiting microbial
CC colonisation, especially of Staphylococcus aureus, in an animal.
CC The combinations can also be used to select donor blood pools for the
CC preparation of purified blood products for passive immunisation.

The present sequence is a serine-aspartate repeat region protein, SdrG from *Staphylococcus epidermidis*. The Sdr protein is useful in vaccine preparation in combination with specific bacterial binding proteins. These vaccines can be used to treat a broad spectrum of bacterial infections, including those arising from both coagulase-positive and coagulase-negative bacteria.

Query Match	20.4%;	Score 353;	DB 21;	Length 991;
Best Local Similarity	30.6%;	Pred. No. 6.1e-17;		
Matches 101;	Conservative	60;	Mismatches 139;	Indels 30;
				Gaps 14;

Qy	13	ITNQLTNVTVGII-DSGTTVYPHQAGYVVKLYGFSVPNSAVKGDTEFKITVPKELNLNGVTS	71
Db	317	VTDQ--SITEGYDDSDGIIKAHDAENLIYDVTFEVDKVKSGDTMTVNIDKNTVPSDLTD	374
Qy	72	TAKVPPIMAGD-QVLANGVID-SDGNVIYTFDYNTKDDVKATLTPAYIDPENVKKTG	129
Db	375	SFAIPKIKDNGSEIIATGYDNTNKQIITYTFTDYVDKYENIKAHLKLTYSIDKSKYPNNN	434
Qy	130	--NVTLATGIGSTTANKTVLVDEYKYGKFYNLSIKGTIDQIDKTNNTYRQTIYVNPSPGD	186
Db	435	TKLDVEYKLTALSS-VNKTIITVEYQKPNENRTANLQSMETNIDTKNHTVEQTIYINPL--	490
Qy	187	NVIAPVLGTNLKPNTDSNALIDQQNTSIKVVYKVDNAADLSESYFV-NPENFEDVTNSVNI	245
Db	491	RYSAKETNVNISGNGDEGSTIIDDSTIIKVKYKVDNQNLPDSPNRIYDYSEYEDVTNDYYA	550
Qy	246	TFPNPNQYKVEFNTPPDDQITTPYIVVVVNGHIDDPNSKGD-----LALRSTLYGYSNII	298
Db	551	QLGNNDNVNIFG---NIDSPYIIKVISKYDPN-KDDYTTIQQTVTMQTTINEXTGE--	603

Qy 299 WRSMSWDNEVAFNNGSGGDGIDKPVVPEQ 328
:**:** :**:** :**:** :**:** :**:** :**:** :**:** :**:** :**:**

Dd 604 FRTASYDNTIAFSTSSGGQG-DLP--PEK 630

RESULT 14	
AAW41602	
ID	AAW41602 standard; Protein; 1092 AA.
XX	
AC	AAW41602;
XX	
DT	22-JUN-1998 (first entry)
XX	
DE	Staphylococcus epidermidis fibrinogen binding protein FIG.
XX	
XX	Fibrinogen binding protein; FIG; aggregation; infection; coagulase-negative Staphylococcus; therapy; diagnosis; immunisation; immunogen; vaccine.

PN WO9748727-A1.
XX
XX
PD 24-DEC-1997.
XX
PF 18-JUN-1997; 97WO-SE01091.

PR 20-JUN-1996; 96SE-0002496.
XX
PA (FLOC/) FLOCK J.
PA (FRYK/) FRYKBERG L.
PA (GUSS/) GUSS B.
PA (LIND/) LINDBERG M.
PA (NILS/) NILSSON M.
XX
PI Flock J, Frykberg L, Guss B, Lindberg M, Nilsson M;
XX
DR WPI; 1998-063079/06.
DR N-PSDB; AAV04279.

Fibrinogen-binding protein from coagulase-negative Staphylococcus -
 PT used for prevention, treatment and diagnosis of Staphylococcus
 PT infection

PS Example 3; Fig 6; 45pp; English.

The protein comprises the fibrinogen binding protein (FIG) of coagulase-negative *Staphylococcus epidermidis* HB. Its amino acid sequence was deduced from the isolated *fig* gene (see AAV04279). The closest known analogue of FIG is the clumping factor of *S. aureus* which also binds fibrinogen and promotes bacterial aggregation in serum. Recombinant FIG polypeptides can be expressed in host cells. They are used as immunogens, particularly in vaccines (which may be expressed *in vivo*) to protect humans and animals against coagulase-negative *Staphylococcus* infection. Antibodies raised against FIG can be used for passive immunisation. They block the adherence of bacteria) and for diagnosis.

SQ Sequence 1092 AA;

Query Match	20.2%	Score 349;	DB 19;	Length 1092;
Best Local Similarity	30.3%;	Pred. No. 1.4e-16;		
Matches 100: Conservative	61;	Mismatches 139;	Indels 30;	

QY	13	ITNQLTNVTVGI-DSGTTVYPHQAGYVVKLNYGFSVPNSAVKGD	TFKITVPKELNLNGVTS	71	
Db	287	VTDQ--SITEGYDDSEGVKAHAENLIYDVTFEVDDKVKSGD	TMTVDIDKNTVPSDLTD	344	
QY	72	TAKVPPIMAGD-QVLANGVIDS-DGNVIYTF	TDYVNTKDDVKATLTMPAYIDPENVKKTG	129	
Db	345	SFTIPKIKDNGSEIIATGYDNKNKQITYT	FTDYVDKYENIKAHLKLTSYIDKSKVPNNN	404	
QY	130	--NVTLATIGIGTTANKTVLVDY	EKYGKFYNLSIKGTIDQIDKTNNTYRQTIYVNP	SGD 186	
Db	405	TKLDVEYKTALSS--VNKTIVEYQRP	NENRTANLQSMFTNIDTKNHTVEQTIYINPL--	460	
QY	187	NVIAPVLTGNLKPNTDSNALIDQONT	SIKVVYKVDNAADLSESYFV-NPENFEDVTNSVNI	245	
Db	461	RYSAKETNVNISNGDEG	STIIDDSTIIKVVYKVGDNQNLPSNR	IYDIYSEYEDVTND	DYA 520
QY	246	TFPNPNQYKVEFNTPD	QITTPYIVVNGHIDPN	SKGD-----LALRSTLYGYN	SNII 298
Db	521	QLGNNDVNINFG----	NIDSPYIIKVISKYDPN-KDDYTTIQQT	VTMQTTINEYTGE--	573
QY	299	WRMSWDNEVAFNNGSGG	DGIDKPVVPEQ	328	
Db	574	FRTASYDNNTIAFSTSSG	GOG-DLP--PEK	600	

RESULT 15
AAU33960
ID AAU33960 standard; Protein: 978 AA.

14-FEB-2002	(first entry)
Staphylococcus aureus	cellular proliferation protein #236.
Antisense;	prokaryotic cellular proliferation protein;

18-JUN-1997; 97WO-SE01091.

PR 23-OCT-2000; 2000US-242578P.
PR 27-NOV-2000; 2000US-253625P.
PR 22-DEC-2000; 2000US-257931P.
PR 16-FEB-2001; 2001US-269308P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
PI Yamamoto RT, Xu HH;
XX
DR WPI; 2001-611495/70.
DR N-PSDB; AAS55104.
XX
PT New polynucleotides for the identification and development of
PT antibiotics, comprise sequences of antisense nucleic acids -
XX
PS Example 3; Seq ID No 12838; 511pp; English.
XX
CC The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the
CC genes, their use in the discovery of novel antibiotics, the essential
CC genes themselves and the encoded proteins. The prokaryotes used are
CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
CC invention is also useful for the identification of potential new targets
CC for antibiotic development. The antisense nucleic acids can also be used
CC to identify proteins used in proliferation, to express these proteins,
CC and to obtain antibodies capable of binding to the expressed proteins.
CC The proteins can be used to screen compounds in rational drug discovery
CC programmes. The antisense nucleic acid sequence is also useful to screen
CC for homologous nucleic acids which are required for cell proliferation in
CC a wide variety of organisms. The present sequence represents an
CC essential prokaryotic cellular proliferation protein.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 1018 AA;

Query Match 19.6%; Score 339; DB 22; Length 1018;
Best Local Similarity 24.9%; Pred. No. 6.5e-16;
Matches 83; Conservative 77; Mismatches 145; Indels 28; Gaps 11;

Qy 10 GTDITNQLTNVTVGIDSG----TTVYPHQAGYVKLNLYGFSVPNSAVKGDTFKITVPKELN 65
Db 194 GTDVTSTAKVPEIKNGSVVMATGEVLEGGKIRYTFNTDIEDKVDVTAELEINLFIDPKTV 252
Db 66 LNGVTSTAKVPPIMAGDQVLANGVIDSDGNVIYTFDYVNTKDDVKATLTMPAYIDPENV 125
Db 253 THGVSTARKVPEIKNGSVVMATGEVLEGGKIRYTFNTDIEDKVDVTAELEINLFIDPKTV 312
Qy 126 KKTGNVTLATGIGSTTANKTVLVVDY-EKYGKFYNLSIKGTIDQIDKTNNTYRQTIYVNP 184
Db 313 QTNGNQITITSTLNEEQTSKELDVYKDGIGNYY-ANLNGSIETFNKANNRFSHVAFIKPN 371
Qy 185 GDNVIAPVLTGNLKPNTDSNALIDQQNTSIKVK-VDNAADLSESYFVN---PENFEDVT 240
Db 372 NGKTSVTVTGTLMKGSNQNG----NQPKVRIFEYLGNNEDIKSVYANTTDTSKFEVT 427
Qy 241 NSV--NITFPNPQYKVEFNTPDDQITTPYIVVVNGHIDPNSKGDALALRSTLYGYSNII 298
Db 428 SNMSGNLNLQNGSYSLNI----ENLDKTYVVHYDGEY-LNGTDEVDFRTQVMVGHPEQLY 482
Qy 299 -----WRSMWDNEVAFNNGSGSGDGIDKPVV 325
Db 483 KYVYDRGYTLTDWNGLVLYSNKANGNGKNGPII 515

RESULT 19
AAP82115
ID AAP82115 standard; protein; 1018 AA.
XX

AC AAP82115;
XX
DT 05-JAN-1990 (first entry)
XX
DE Fibronectin binding protein.
XX
KW Fibronectin binding protein; S.aureus; immunisation;
KW mastitis; ruminants; wound infection; diagnosis.
XX
OS Staphylococcus aureus.
XX
PN EP294349-A.
XX
PD 07-DEC-1988.
XX
PF 30-MAY-1988; 88EP-0850188.
XX
PR 01-JUN-1987; 87SE-0002272.
XX
PA (ALFA) ALFA-LAVAL AGRI INTERNATIONAL AB.
XX
PI Hook M, Lindberg MK, Signas LC, Wadstrom TM, Froman G;
XX
DR WPI; 1988-347978/49.
DR P-PSDB; AAP82115.
XX
PT Hybrid DNA encoding Staphylococcus aureus fibronectin binding protein -
PT useful for immunisation and topical application to prevent
PT staphylococcal infections.
XX
PS Disclosure; Fig. 8A:1-8A:2; 23pp; English.
XX
CC The S.aureus fibronectin binding protein may be included in a hybrid
CC protein. The protein may be used to immunise ruminants against
CC staphylococcal mastitis, pref. when given at 0.5-5 microg/kg, using 3
CC doses at 1-3 week intervals; and for topical applicn. to prevent wound
CC infection, pref. using isotonic saline soln. of concn. 25-250 microg/mg.
CC The DNA encoding the binding protein and the hybrid protein may be
CC immobilised on a carrier and used to diagnose staphylococcal infections.
XX
SQ Sequence 1018 AA;

Query Match 19.1%; Score 331; DB 9; Length 1018;
Best Local Similarity 24.6%; Pred. No. 2.4e-15;
Matches 82; Conservative 77; Mismatches 146; Indels 28; Gaps 11;

Qy 10 GTDITNQLTNVTVGIDSG----TTVYPHQAGYVKLNLYGFSVPNSAVKGDTFKITVPKELN 65
Db 194 GTDVTSTAKVPEIKNGSVVMATGEVLEGGKIRYTFNTDIEDKVDVTAELEINLFIDPKTV 252
Qy 66 LNGVTSTAKVPPIMAGDQVLANGVIDSDGNVIYTFDYVNTKDDVKATLTMPAYIDPENV 125
Db 253 THGVSTARKVPEIKNGSVVMATGEVLEGGKIRYTFNTDIEDKVDVTAELEINLFIDPKTV 312
Qy 126 KKTGNVTLATGIGSTTANKTVLVVDY-EKYGKFYNLSIKGTIDQIDKTNNTYRQTIYVNP 184
Db 313 QTNGNQITITSTLNEEQTSKELDVYKDGIGNYY-ANLNGSIETFNKANNRFSHVAFIKPN 371
Qy 185 GDNVIAPVLTGNLKPNTDSNALIDQQNTSIKVK-VDNAADLSESYFVN---PENFEDVT 240
Db 372 NGKTSVTVTGTLMKGSNQNG----NQPKVRIFEYLGNNEDIKSVYANTTDTSKFEVT 427
Qy 241 NSV--NITFPNPQYKVEFNTPDDQITTPYIVVVNGHIDPNSKGDALALRSTLYGYSNII 298
Db 428 SNMSGNLNLQNGSYSLNI----ENLDKTYVVHYDGEY-LNGTDEVDFRTQVMVGHPEQLY 482
Qy 299 -----WRSMWDNEVAFNNGSGSGDGIDKPVV 325
Db 483 KYVYDRGYTLTDWNGLVLYSNKANGNGKNGPII 515

RESULT 20
AAW89806

CC may be used to coat a medical device to reduce the S. aureus infection of
CC an indwelling medical device, especially where the medical device is
CC selected from the group consisting of vascular grafts, vascular stents,
CC intravenous catheters, artificial heart valves, and cardiac assist
CC devices. ClfB, SdrC, SdrD, SdrE, or an active fragment, subdomain or
CC encoding gene may be used as a vaccine. The DS (aspartate serine) repeat
CC region or a gene encoding it may be used as an identifying probe for the
CC identification of genes and encoding proteins from Staphylococcus aureus
CC (other than ClfA), S. hemolyticus, S. lugdenensis, and S. schleriferi.
*CC The proteins of the invention have antibacterial activity.
XX
SQ Sequence 1166 AA;

Query Match 18.7%; Score 323.5; DB 20; Length 1166;
Best Local Similarity 31.2%; Pred. No. 1e-14;
Matches 111; Conservative 53; Mismatches 137; Indels 55; Gaps 18;
QY 2 VAADAPAAAGTDITNQLTNVT-----VGIDSGTTVYPHOAGYVKLNYGFSVPNSAVKGDTF 56
II I I : : I I I I I I : : : : I I I I
269 VAQPAAVASNNV-NDLIKVTKQTIKVGDKDNVAAAHDKDIEYDTEFTIDNKVKKGDTM 327
57 KI-----TVPKELNLNGVTSTAKVPPIMAGD---QVLANGVID-SDGNVIYFTFDYVNTK 107
I I : I I I I I I I I I I I I : : : : I I I I I I
328 TINYDKNVPSDL-----TDKNPDIDITDPSGEVIAKGTDFDKATKQIITYFTFDYVDKY 380
108 DDVKATLTMPAYIDPENV--KKTGNVTLATGIGSTTANKTVLVLDYKYGKFFYNLSKGTI 165
: : : : I I : : I I I I I I : : I I I I : : : :
381 EDIKSRLTLYSIDKKTVPNETSLNLTFTAGKETSON--VTVDYQDPMVHGDSNIQSIF 438
166 DQIDKTNNTYRQTIYVNP---SGDNVIAPVL-----TGNLKPNTDSNALIDQNTSIKV 216
: : : : I I I I I I I I : : I I I I : : I I I I I I
439 TKLDEDKQTIEQQIYVNPPLKKSATNTKVDIAGSQVDDYGNIKLNGS-TIID-QNTEIKV 496
217 YKVDNAADLSESYFV-NPENFEDVTSNVNITFPNPNQYKVEFNTPD-DQITTPYIVVVNG 274
I I I I : : : : I I I I : : I I : : I I I I I I
497 YKVNDSQQLPQSNRIYDFSOYEDVTSQ----FDNKKFSNNVATLDFGDINSAYIIKVVVS 552
275 HIDPNSKGDG-----ALRST-LYGYSNIIWRMSWDNEVAFNNGSGSGDGIDKP 323
I I I I : : : : I I I I : : : : I I I I I I
553 KYTPTSDGELDIAQGTSMRTDKYGY-----YNYAGYSNFIVTSNDTGGGDGTVKP 603

RESULT 22
AAAY08640
ID AAY08640 standard; protein; 918 AA.
XX
AC AAY08640;
09-AUG-1999 (first entry)
DE S. aureus ClfB protein.
XX
KW Fibrinogen-binding protein; alpha chain; beta chain; ClfB; SdrC; SdrD;
KW SdrE; fibrinogen; medical device; competitive inhibitor; pharmaceutical;
KW treatment; infection; septicemia; osteomyelitis; mastitis; endocarditis;
KW extracellular matrix; vascular graft; vascular stent; vaccine;
KW intravenous catheter; artificial heart valve; cardiac assist device;
KW antibacterial.
XX
OS Staphylococcus aureus.
XX
PN WO9927109-A2.
XX
PD 03-JUN-1999.
XX
PF 25-NOV-1998; 98WO-US25246.
XX
PR 31-AUG-1998; 98US-0098427.
PR 26-NOV-1997; 97US-0066815.
XX
PA (EIDH/) EIDHIN D N.
PA (FORF-) FORFAS T/A BIORESEARCH IRELAND.
PA (FOST/) FOSTER T J.

(HOOK/) HOOK M A O.
(INH-) INHIBITEX INC.
(JOSE/) JOSEFSSON E.
(PATT/) PATTI J M.
(PERK/) PERKINS S E.
XX EIDHIN DN, Foster TJ, Hook MAO, Josefsson E, Patti JM;
PI Perkins SE;
XX WPI; 1999-357844/30.
DR N-PSDB; AAX77591.
XX
PT Staphylococcus aureus fibrinogen-binding proteins for treating
PT septicemia, osteomyelitis, mastitis or endocarditis
XX
PS Claim 2; Fig 5; 143pp; English.
XX
CC This invention describes novel Staphylococcus aureus fibrinogen-binding
CC proteins that bind both the alpha and beta fibrinogen chains. The
CC proteins (and their encoding nucleic acids are ClfB, SdrC, SdrD and
CC SdrE). Staphylococcus aureus is thought to utilize fibrinogen to adhere
CC to medical devices, binding proteins that bind both the alpha and beta
CC fibrinogen chains (ClfB, SdrC, SdrD and SdrE) can therefore be used as
CC competitive inhibitors to block this binding. Antibodies against ClfB,
CC SdrC, SdrD and SdrE inhibit ClfB, SdrC, SdrD and SdrE mediated binding.
CC The proteins of the invention can be used in a pharmaceutical composition
CC for the treatment of Staphylococcus aureus infection e.g. septicemia,
CC osteomyelitis, mastitis or endocarditis or to inhibit the binding of
CC S. aureus to the extracellular matrix. The proteins or their fragments
CC may be used to coat a medical device to reduce the S. aureus infection of
CC an indwelling medical device, especially where the medical device is
CC selected from the group consisting of vascular grafts, vascular stents,
CC intravenous catheters, artificial heart valves, and cardiac assist
CC devices. ClfB, SdrC, SdrD, SdrE, or an active fragment, subdomain or
CC encoding gene may be used as a vaccine. The DS (aspartate serine) repeat
CC region or a gene encoding it may be used as an identifying probe for the
CC identification of genes and encoding proteins from Staphylococcus aureus
CC (other than ClfA), S. hemolyticus, S. lugdenensis, and S. schleriferi.
CC The proteins of the invention have antibacterial activity.
XX
SQ Sequence 918 AA;

Query Match 18.0%; Score 312.5; DB 20; Length 918;
Best Local Similarity 24.9%; Pred. No. 4.5e-14;
Matches 92; Conservative 69; Mismatches 132; Indels 77; Gaps 15;
QY 6 APAAGTDITNQLTNVTGIDSGTTVYPHOAGYVKLNYGFSVPNSAVKGTFTKITVPKELN 65
I I I I : : : : I I I I I I : : I I I I
213 ADAKGTNVNDKVTASNEKLEK-TTFDPNQSGNTFMAANFTVTDKVKSGDYF----- 262
66 LNGVTSTAKVPPIMAGDQVLANGVID-----SDGNVI----- 97
263 -----TAKLP-----DSLTCNGDGDVYSNSNNTMPIADIKSTNGDVVAKATYDILTKTYT 311
98 YFTFDYVNTKDDVKATLTMPAYIDPENVKKTGNVTLATGIGSTTANKTVLVLDYK-----Y 153
: I I I I I I : : : : I I I I I I : : I I I I
312 FVFTDYVNNKENINGQFSLPLFTDRAKAPKSGTYDANINIADEMFENNKITYNYSSPIAGI 371
154 GKEYNLSIKGTIDQIDKTN--NTYRQTIYVNPS-----GDNVIAPLVTGNLKPNTDSNALI 207
I I I I : : I I : I I I I I I I I : : : : I I : :
372 DKPNGANISSQIIGVDTASGQNTYKQTVFVNPVKQVRLGNTWV--YIKGYQDKIESSGKV 429
208 DQONTSIKVYKVDNAADLSESYFVNP--ENFEDVTSNV--NITFPNPNQYKKEFNTPDDQ 263
: I : : : : : : : : : : : : : : : : : : : : : : : : : :
430 SATDTKLRIFEVNDTSKLSDSYADPNDSNLKEVTDQFKNRIYEHHPNVASIKFG-----D 485
264 ITTPYIVVNVNGHIDPNSKGLALRSTLYGYN----SNIWRMSWDNEVAFNNGSGSGDG 319
I I I I I I I I I I : : : : I I I I I I I I I I
486 ITKTYVVLVEGHYDNTGKN---LKTQVIQENVDPVTNRDYSIFGWNNEVNVRYGGGSADG 542
320 IDKPVVPEQP 329
I I I I : :
543 -DSAVNPKDP 551

CC may be used to coat a medical device to reduce the S. aureus infection of
CC an indwelling medical device, especially where the medical device is
CC selected from the group consisting of vascular grafts, vascular stents,
CC intravenous catheters, artificial heart valves, and cardiac assist
CC devices. ClfB, SdrC, SdrD, SdrE, or an active fragment, subdomain or
CC encoding gene may be used as a vaccine. The DS (aspartate serine) repeat
CC region or a gene encoding it may be used as an identifying probe for the
CC identification of genes and encoding proteins from Staphylococcus aureus
CC (other than Cifa), S. hemolyticus, S. lugdenensis, and S. schleriferi.
CC The proteins of the invention have antibacterial activity.

XX
SQ Sequence 1315 AA;
Query Match 14.9%; Score 257.5; DB 20; Length 1315;
Best Local Similarity 26.9%; Pred. No. 6.5e-10;
Matches 88; Conservative 51; Mismatches 157; Indels 31; Gaps 12;
QY 13 ITNQLTNVTVGIDSGTTVYPHQAGYVKLVNGYFSPNSAVKGDFTFKITVPKELNLNGVT-- 70
251 ITSNTTLTVVDADKNNKIVPAQ-DYLSLKSQITVDDKVKSGDYFTIKYSDTVQVYGLNPE 309
71 -----STAKVPPIMAGDQVLANGVIDSDGNVI-YTFTDYVNTKDDVKATLTMPAYIDPEN 124
Db 310 DIKNIGIDKP---NNGETIATAKHDTANNLITYTFTDYVDRFNSVQMGINYSIYMDADT 366
QY 125 VKKTGN-VTLATGIGSTTANKTTLVDYVEKYGKFYNLSIKG-----TIDQIDKTNNT--YRQ 177
367 IPVSKNDVEFNVTIGNTTTKTATTANIQYPDYVVVNEKNSIGSAFTETVSHVGNKENPGYKQ 426
QY 178 TIYVNPSEGNVIAPIVLTGNLKPNTDSNAL--IDQNTSIKVKYVDNAADLSESYFVNPN 235
Db 427 TIYVNPSESLTNAKLKVQAYHSSYPNNIGQINKDVTDIKIYQVPKGYTLNKGVDVNTKE 486
QY 236 FEDVTNSV--NITFPNPQYKVEFNTPDDQITTPYIVVNVNGHID-PNSKGDALRSTLYG 292
Db 487 LTDVTNQYLQKITYGDNNSAVIDFGNAD----SAYVVMVNTKFOYTNSPTLVQMATLS 542
QY 293 YNSNIWRSMWDNEVAFNNGSGGDG 319
Db 543 STGN---KSVSTGNALGFTNNQSGGAG 566

RESULT 25
AAU34402
ID AAU34402 standard; Protein; 1349 AA.
XX
AC AAU34402;
14-FEB-2002 (first entry)
DE Staphylococcus aureus cellular proliferation protein #678.
XX
KW Antisense; prokaryotic cellular proliferation protein;
KW antibiotic; antibacterial; drug design.
XX
OS Staphylococcus aureus.
XX
PN WO200170955-A2.
XX
PD 27-SEP-2001.
XX
PF 21-MAR-2001; 2001WO-US09180.
XX
PR 21-MAR-2000; 2000US-191078P.
PR 23-MAY-2000; 2000US-206848P.
PR 26-MAY-2000; 2000US-207727P.
PR 23-OCT-2000; 2000US-242578P.
PR 27-NOV-2000; 2000US-253625P.
PR 22-DEC-2000; 2000US-257931P.
PR 16-FEB-2001; 2001US-269308P.
XX
PA (ELIT-) ELITRA PHARM INC.

PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
PI Yamamoto RT, Xu HH;
XX
DR WPI; 2001-611495/70.
DR N-PSDB; AAS52261.
XX
PT New polynucleotides for the identification and development of
PT antibiotics, comprise sequences of antisense nucleic acids -
XX
PS Example 3; Seq ID No 5898; 511pp; English.
XX
CC The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the
CC genes, their use in the discovery of novel antibiotics, the essential
CC genes themselves and the encoded proteins. The prokaryotes used are
CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
CC invention is also useful for the identification of potential new targets
CC for antibiotic development. The antisense nucleic acids can also be used
CC to identify proteins used in proliferation, to express these proteins,
CC and to obtain antibodies capable of binding to the expressed proteins.
CC The proteins can be used to screen compounds in rational drug discovery
CC programmes. The antisense nucleic acid sequence is also useful to screen
CC for homologous nucleic acids which are required for cell proliferation in
CC a wide variety of organisms. The present sequence represents an
CC essential prokaryotic cellular proliferation protein.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

XX
SQ Sequence 1349 AA;
Query Match 14.9%; Score 257.5; DB 22; Length 1349;
Best Local Similarity 26.9%; Pred. No. 6.7e-10;
Matches 88; Conservative 51; Mismatches 157; Indels 31; Gaps 12;
QY 13 ITNQLTNVTVGIDSGTTVYPHQAGYVKLVNGYFSPNSAVKGDFTFKITVPKELNLNGVT-- 70
Db 251 ITSNTTLTVVDADKNNKIVPAQ-DYLSLKSQITVDDKVKSGDYFTIKYSDTVQVYGLNPE 309
QY 71 -----STAKVPPIMAGDQVLANGVIDSDGNVI-YTFTDYVNTKDDVKATLTMPAYIDPEN 124
Db 310 DIKNIGIDKP---NNGETIATAKHDTANNLITYTFTDYVDRFNSVQMGINYSIYMDADT 366
QY 125 VKKTGN-VTLATGIGSTTANKTTLVDYVEKYGKFYNLSIKG-----TIDQIDKTNNT--YRQ 177
Db 367 IPVSKNDVEFNVTIGNTTTKTATTANIQYPDYVVVNEKNSIGSAFTETVSHVGNKENPGYKQ 426
QY 178 TIYVNPSEGNVIAPIVLTGNLKPNTDSNAL--IDQNTSIKVKYVDNAADLSESYFVNPN 235
Db 427 TIYVNPSESLTNAKLKVQAYHSSYPNNIGQINKDVTDIKIYQVPKGYTLNKGVDVNTKE 486
QY 236 FEDVTNSV--NITFPNPQYKVEFNTPDDQITTPYIVVNVNGHID-PNSKGDALRSTLYG 292
Db 487 LTDVTNQYLQKITYGDNNSAVIDFGNAD----SAYVVMVNTKFOYTNSPTLVQMATLS 542
QY 293 YNSNIWRSMWDNEVAFNNGSGGDG 319
Db 543 STGN---KSVSTGNALGFTNNQSGGAG 566

Search completed: June 23, 2003, 14:01:38
Job time : 56.406 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 23, 2003, 14:00:06 ; Search time 19.367 Seconds
(without alignments)
502.864 Million cell updates/sec

Title: US-10-056-052A-4
Perfect score: 1732
Sequence: 1 MVAADAPAAAGTDITNQLTNV.....NGSGSGDIDKPVVPEQPDE 331

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues 262574
Number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1732	100.0	345	4	US-08-856-253-7	Sequence 7, Appli
2	1727	99.7	933	3	US-08-293-728-2	Sequence 2, Appli
3	1727	99.7	933	4	US-09-421-868-2	Sequence 2, Appli
4	353	20.4	930	4	US-09-134-001C-5314	Sequence 5314, Ap
5	204	11.8	1112	2	US-08-714-402-2	Sequence 2, Appli
6	204	11.8	1161	4	US-09-327-536-2	Sequence 2, Appli
7	121.5	7.0	992	4	US-09-206-942-61	Sequence 61, Appl
8	121.5	7.0	998	4	US-09-206-942-59	Sequence 59, Appl
9	120	6.9	737	4	US-09-071-035-460	Sequence 460, App
10	120	6.9	2032	4	US-09-071-035-458	Sequence 458, App
11	120	6.9	2032	4	US-09-071-035-462	Sequence 462, App
12	120	6.9	2032	4	US-09-071-035-466	Sequence 466, App
13	118	6.8	789	4	US-09-002-285-94	Sequence 94, Appl
14	117	6.8	1180	4	US-09-206-942-65	Sequence 65, Appl
15	117	6.8	1188	4	US-09-206-942-63	Sequence 63, Appl
16	116	6.7	789	4	US-09-002-285-96	Sequence 96, Appl
17	115	6.6	790	4	US-08-960-780-4	Sequence 4, Appli
18	115	6.6	790	4	US-09-073-898-4	Sequence 29, Appl
19	114	6.6	789	1	US-08-471-033-29	Sequence 29, Appl
20	114	6.6	789	2	US-08-471-044-29	Sequence 29, Appl
21	114	6.6	789	2	US-08-463-483A-29	Sequence 29, Appl
22	114	6.6	789	2	US-08-471-046A-29	Sequence 29, Appl
23	114	6.6	789	2	US-08-470-566B-29	Sequence 29, Appl
24	114	6.6	789	2	US-08-838-219B-2	Sequence 2, Appli
25	114	6.6	789	2	US-08-469-334-29	Sequence 29, Appl
26	114	6.6	789	3	US-09-300-529-29	Sequence 29, Appl
27	114	6.6	789	3	US-09-233-336A-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-08-856-253-7
; Sequence 7, Application US/08856253
; Patent No. 6288214
; GENERAL INFORMATION:
; APPLICANT: Hook, Magnus
; APPLICANT: Patti, Joseph M.
; APPLICANT: House-Pompeo, Karen
; APPLICANT: Sthanam, Narayana
; APPLICANT: Symersky, Jindrich
; TITLE OF INVENTION: COLLAGEN BINDING PROTEIN COMPOSITIONS
; TITLE OF INVENTION: AND METHODS OF USE
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: U.S.
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/856.253
; FILING DATE: Concurrently Herewith
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/017,678
; FILING DATE: 16-MAY-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Kitchell, Barbara S.
; REGISTRATION NUMBER: 33,928
; REFERENCE/DOCKET NUMBER: TAMK:193
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (512) 418-3000
; TELEFAX: (512) 474-7577
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 345 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
US-08-856-253-7

Query Match 100.0%; Score 1732; DB 4; Length 345;
Best Local Similarity 100.0%; Pred. No. 6e-143;
Matches 331; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVAADAPAAAGTDITNQLTNVTGIDSGTTVYPHQAGYVKLVNYSVPNSAVKGDTEKITY 60
Db 13 MVAADAPAAAGTDITNQLTNVTGIDSGTTVYPHQAGYVKLVNYSVPNSAVKGDTEKITY 72
QY 61 PKELNLNGVTSTAKVPPIMAGDQVLANGVIDSDGNVIYTFDYVNTKDDVKATLTMPAYI 120
Db 73 PKELNLNGVTSTAKVPPIMAGDQVLANGVIDSDGNVIYTFDYVNTKDDVKATLTMPAYI 132
QY 121 DPENVKKTGNVTLATGIGSTTANKTVLVVDYKYGKFNLSIKGTIDQIDKTNNTYRQTIY 180
Db 133 DPENVKKTGNVTLATGIGSTTANKTVLVVDYKYGKFNLSIKGTIDQIDKTNNTYRQTIY 192
QY 181 VNPSGDNVIAPVLTGNLKPNTDSNALIDQQNTSIKVKYKVDNAADLSESYFVNPFEDVT 240
Db 193 VNPSGDNVIAPVLTGNLKPNTDSNALIDQQNTSIKVKYKVDNAADLSESYFVNPFEDVT 252
QY 241 NSVNITFPNPNOYKVEFNTPDQITTPYIVVNGHIDPNKSGDLALRSTLYGYSNIIWR 300

Db 253 NSVNITFPNPNOYKVEFNTPDQITTPYIVVNGHIDPNKSGDLALRSTLYGYSNIIWR 312
QY 301 SMSWDNEVAFNNGSGDGDIDKPKVPVPEQPD 331
Db 313 SMSWDNEVAFNNGSGDGDIDKPKVPVPEQPD 343

RESULT 2
US-08-293-728-2
; Sequence 2, Application US/08293728D
; Patent No. 6008341
; GENERAL INFORMATION:
; APPLICANT: Foster, Timothy J.
; APPLICANT: McDevitt, Damien L.
; TITLE OF INVENTION: The S. aureus Fibrinogen Binding Protein Gene
; FILE REFERENCE: 05344.105011
; CURRENT APPLICATION NUMBER: US/08/293,728D
; CURRENT FILING DATE: 1994-08-22
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 933
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-08-293-728-2

Query Match 99.7%; Score 1727; DB 3; Length 933;
Best Local Similarity 100.0%; Pred. No. 7.1e-142;
Matches 330; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAADAPAAAGTDITNQLTNVTGIDSGTTVYPHQAGYVKLVNYSVPNSAVKGDTEKITY 61
Db 221 VAADAPAAAGTDITNQLTNVTGIDSGTTVYPHQAGYVKLVNYSVPNSAVKGDTEKITY 280
QY 62 KELNLNGVTSTAKVPPIMAGDQVLANGVIDSDGNVIYTFDYVNTKDDVKATLTMPAYI 121
Db 281 KELNLNGVTSTAKVPPIMAGDQVLANGVIDSDGNVIYTFDYVNTKDDVKATLTMPAYI 340
QY 122 PENVKKTGNVTLATGIGSTTANKTVLVVDYKYGKFNLSIKGTIDQIDKTNNTYRQTIY 181
Db 341 PENVKKTGNVTLATGIGSTTANKTVLVVDYKYGKFNLSIKGTIDQIDKTNNTYRQTIY 400
QY 182 NPNGDNVIAPVLTGNLKPNTDSNALIDQQNTSIKVKYKVDNAADLSESYFVNPFEDVTN 241
Db 401 NPNGDNVIAPVLTGNLKPNTDSNALIDQQNTSIKVKYKVDNAADLSESYFVNPFEDVTN 460
QY 242 SVNITFPNPNOYKVEFNTPDQITTPYIVVNGHIDPNKSGDLALRSTLYGYSNIIWR 301
Db 461 SVNITFPNPNOYKVEFNTPDQITTPYIVVNGHIDPNKSGDLALRSTLYGYSNIIWR 520
QY 302 MSWDNEVAFNNGSGDGDIDKPKVPVPEQPD 331
Db 521 MSWDNEVAFNNGSGDGDIDKPKVPVPEQPD 550

RESULT 3
US-09-421-868-2
; Sequence 2, Application US/09421868
; Patent No. 6177084
; GENERAL INFORMATION:
; APPLICANT: Foster, Timothy J.
; APPLICANT: McDevitt, Damien L.
; TITLE OF INVENTION: The S. aureus Fibrinogen Binding Protein Gene
; FILE REFERENCE: 05344.105011
; CURRENT APPLICATION NUMBER: US/09/421,868
; CURRENT FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 08/293,728
; PRIOR FILING DATE: 1994-08-22
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 933

TYPE: PRT
ORGANISM: Staphylococcus aureus
US-09-421-868-2
Query Match 99.7%; Score 1727; DB 4; Length 933;
Best Local Similarity 100.0%; Pred. No. 7.1e-142;
Matches 330; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 2 VAADAPAAAGTDITNQLTNVTGIDSGTTVYPHQAGYVVKLVNFGFVSPNSAVKGDTEKITVP 61
Db 221 VAADAPAAAGTDITNQLTNVTGIDSGTTVYPHQAGYVVKLVNFGFVSPNSAVKGDTEKITVP 280
Qy 62 KELNLNGVTSTAKVPPIMAGDQVLANGVIDSDGNVIYTFDYVNTKDDVKATLTMPAYID 121
Db 281 KELNLNGVTSTAKVPPIMAGDQVLANGVIDSDGNVIYTFDYVNTKDDVKATLTMPAYID 340
Qy 122 PENVKKTGNVTLATGIGSTTANKTVLVVDYKYGKFNLSIKGTIDQIDKTNNTYRQTIYV 181
Db 341 PENVKKTGNVTLATGIGSTTANKTVLVVDYKYGKFNLSIKGTIDQIDKTNNTYRQTIYV 400
182 NPSGDNVIAPVLTGNLKPNTDSNALIDQONTSIKVYKVDNAADLSESYFVNPFEDVTN 241
Db 401 NPSGDNVIAPVLTGNLKPNTDSNALIDQONTSIKVYKVDNAADLSESYFVNPFEDVTN 460
Qy 242 SVNITFPNPNOYKVEFTPPDDQITTPYIVVNGHIDPNKSGDLALRSTLYGYNLSNIWRS 301
Db 461 SVNITFPNPNOYKVEFTPPDDQITTPYIVVNGHIDPNKSGDLALRSTLYGYNLSNIWRS 520
Qy 302 MSWDNEVAFNNGSGDGDIDKPVVPEQDPE 331
Db 521 MSWDNEVAFNNGSGDGDIDKPVVPEQDPE 550
RESULT 4
US-09-134-001C-5314
Sequence 5314, Application US/09134001C
Patent No. 6380370
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: GTC-007
CURRENT APPLICATION NUMBER: US/09/134,001C
CURRENT FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/064,964
PRIOR FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: US 60/055,779
PRIOR FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 5674
SEQ ID NO 5314
LENGTH: 930
TYPE: PRT
ORGANISM: Staphylococcus epidermidis
US-09-134-001C-5314
Query Match 20.4%; Score 353; DB 4; Length 930;
Best Local Similarity 30.6%; Pred. No. 2.7e-22;
Matches 101; Conservative 60; Mismatches 139; Indels 30; Gaps 14;
Qy 13 ITNQLTNVTVGI-DSGTTVYPHQAGYVVKLVNFGFVSPNSAVKGDTEKITVPKELNLNGVTS 71
Db 284 VTDQ--SITEGYDDSDGIIKAHAENLIYDVTFEVDKVKSGDTMTVNIDKNTVPSDLTD 341
Qy 72 TAKVPPIMAGD-QVLANGVID-SDGNVIYTFDYVNTKDDVKATLTMPAYIDPENVKKTG 129
Db 342 SFAIPKIKDSGEIIATGTYDNTNKQITTYTFDYVDKYENKAHLKLTSDYDKSKVPNNN 401
Qy 130 ---NVTLATGIGSTTANKTVLVVDYKYGKFNLSIKGTIDQIDKTNNTYRQTIYVNPSPGD 186
Db 402 TKLDVEYKTALSS--VNKTTVEYQKPNENRTANLQSMFTNIDTKNHTVEQTIYINPL-- 457
Qy 187 NVIAPVLTGNLKPNTDSNALIDQONTSIKVYKVDNAADLSESYFV-NPENFEDVTNSVNI 245

Db 458 RYSAKETNVNISNGDEGSTIIDSTIIKVYKVGDNQNLPSNRIYDYSEYEDVTNDDYA 517
Qy 246 TFPNPNOYKVEFTPPDDQITTPYIVVNGHIDPNKSGD-----LALRSTLYGYNLSII 298
Db 518 QLGNNNDVNINFG----NIDSPYIIKVISKYDPN-KDDYTTIQOTVTMQTTINEYTG-- 570
Qy 299 WRSMSWDNEVAFNNGSGDGDIDKPVVPEQ 328
Db 571 FRTASYDNTIAFSTSSGQGG-DLP--PEK 597
RESULT 5
US-08-714-402-2
Sequence 2, Application US/08714402
Patent No. 5910441
GENERAL INFORMATION:
APPLICANT: ROCHA, Claudia
APPLICANT: FISCHETTI, Vincent A.
TITLE OF INVENTION: FIBRONECTIN AND FIBRINOGEN BINDING
TITLE OF INVENTION: PROTEIN FROM GROUP A STREPTOCOCCI
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/714,402
FILING DATE: 16-SEP-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 016921-097
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1112 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-714-402-2

Query Match 11.8%; Score 204; DB 2; Length 1112;
Best Local Similarity 22.8%; Pred. No. 3.3e-09;
Matches 86; Conservative 72; Mismatches 145; Indels 74; Gaps 18;
Qy 8 AAGTDTN--QLTNVTVGIDS--GTT-----VYPHQAGYVVKLVNFGFVSPNSAV---K 52
Db 142 AGSKDVSSSLQLENPKMSVSVKYGKTEVSSGAADFYNHAAFKMSFELKQDKSETINP 201
Qy 53 GDFEKITVPKELNLNGVTSTAKVPPIM--AGDQVLANGVIDSDGN-VIYTFDYVNTKDD 109
Db 202 GDTFVLQDRLRLNPKGISQ--DIPKIIYDSANSPLAIGKYHAENHQLIYTFDYIAGLDK 259
Qy 110 VKATLTMPAYIDPE-----NVKKT-----GNVTLATGIGSTTANKTVLVVDYE 151
Db 260 VQLSAELSLFLENKEVLENTSISNFKSTIGGQEITYKGTNVNLYGNESTKESNYITNGLS 319
Qy 152 KYGKFNLSIKGTIDQIDKTNNTYRQTIYVNPSPGDNVIAPVLTGNL----KPNTDSNALI 207
Db 320 NVG-----GSIESYNTETGEFVWYVYVNPRTNI--PYATMNLWGFGGRARSNTSDLE 369

Db 204 AGPTDISDYFTGDETTIIDNFEDPIYLNPDGTPATPPYKEDVTIHWNFNWSIPEDVREQM 263
QY 53 --GDTEKITVPKELNLNGVTSTAKVPPIMAGDQVLANG-----VIDSDGNVIYFTDYV 104
Db 264 KAGDYFEFQLPGNLKPNKPGS-----GDLVDAEGNVYGTYTISEDGTVRFTENERI 314
QY 105 NTKDDVKATLTMPAYIDPENVKKTGNVTLATGIGSTTANKTVLVLDYKYGKFYNLSIKGT 164
Db 315 TSESDIHGDFSLDTHLNDSDGRGPGDWI--DIPTQEDLPPVVIPI-----VPDT 362
QY 165 IDQIDKTNTYRQTIYVNPSPGDNVIAVLGTGNLKPNTDSNALIDQQN----- 211
Db 363 EQQIDKQGHFDRTP--NPS-----AITWTVDIN---QAMKDQTNPTVTETWPTGNTF 409
QY 212 TSIKVKYKVDNAAD--LSE-SYFVNPENFE-DVTNSVNIITFPNPQYKVEFNTPDQDQITTP 267
Db 410 KSVKVELVMNLDGTIKEVGRELSPDEYTVDKNGNVTIKGDTNKAYRLEYQTTIDEAVIP 469
QY 268 ---YIVVVNGHI---DPNSKGDALRSTLYGYNSNIIRWSMSWDN-----EVAFNNG 313
Db 470 DGGGDVPFKNHATLTSDNNPNGLDAEATVTATYGMKMLDKRNDIDYDEANQEFWEINYNVG 529

RESULT 11
US-09-071-035-462
; Sequence 462, Application US/09071035
; Patent No. 6448043
; GENERAL INFORMATION:
; APPLICANT: Gil H. Choi
; TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides
; NUMBER OF SEQUENCES: 496
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/071,035
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: A. Anders Brookes
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB369P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 462:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2032 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-071-035-462

Query Match 6.9%; Score 120; DB 4; Length 2032;
Best Local Similarity 21.9%; Pred. No. 0.17;
Matches 79; Conservative 54; Mismatches 139; Indels 88; Gaps 18;
QY 8 AAGTDITNQLT-NVTVGIDS-----GTTVYPHQAGYVKL--NYGFSVPNSAVK-- 52
Db 204 AGPTDISDYFTGDETTIIDNFEDPIYLNPDGTPATPPYKEDVTIHWNFNWSIPEDVREQM 263

Db 204 AGPTDISDYFTGDETTIIDNFEDPIYLNPDGTPATPPYKEDVTIHWNFNWSIPEDVREQM 263
QY 53 --GDTEKITVPKELNLNGVTSTAKVPPIMAGDQVLANG-----VIDSDGNVIYFTDYV 104
Db 264 KAGDYFEFQLPGNLKPNKPGS-----GDLVDAEGNVYGTYTISEDGTVRFTENERI 314
QY 105 NTKDDVKATLTMPAYIDPENVKKTGNVTLATGIGSTTANKTVLVLDYKYGKFYNLSIKGT 164
Db 315 TSESDIHGDFSLDTHLNDSDGRGPGDWI--DIPTQEDLPPVVIPI-----VPDT 362
QY 165 IDQIDKTNTYRQTIYVNPSPGDNVIAVLGTGNLKPNTDSNALIDQQN----- 211
Db 363 EQQIDKQGHFDRTP--NPS-----AITWTVDIN---QAMKDQTNPTVTETWPTGNTF 409
QY 212 TSIKVKYKVDNAAD--LSE-SYFVNPENFE-DVTNSVNIITFPNPQYKVEFNTPDQDQITTP 267
Db 410 KSVKVELVMNLDGTIKEVGRELSPDEYTVDKNGNVTIKGDTNKAYRLEYQTTIDEAVIP 469
QY 268 ---YIVVVNGHI---DPNSKGDALRSTLYGYNSNIIRWSMSWDN-----EVAFNNG 313
Db 470 DGGGDVPFKNHATLTSDNNPNGLDAEATVTATYGMKMLDKRNDIDYDEANQEFWEINYNVG 529

RESULT 12
US-09-071-035-466
; Sequence 466, Application US/09071035
; Patent No. 6448043
; GENERAL INFORMATION:
; APPLICANT: Gil H. Choi
; TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides
; NUMBER OF SEQUENCES: 496
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/071,035
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: A. Anders Brookes
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB369P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 466:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2032 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-071-035-466

Query Match 6.9%; Score 120; DB 4; Length 2032;
Best Local Similarity 21.9%; Pred. No. 0.17;
Matches 79; Conservative 54; Mismatches 139; Indels 88; Gaps 18;
QY 8 AAGTDITNQLT-NVTVGIDS-----GTTVYPHQAGYVKL--NYGFSVPNSAVK-- 52
Db 204 AGPTDISDYFTGDETTIIDNFEDPIYLNPDGTPATPPYKEDVTIHWNFNWSIPEDVREQM 263

Qy 53 --GDFPKITVPKELNLNGVTSTAKVPPIMAGDQVLANG-----VIDSDGNVIYTFDTYV 104
Db 264 KAGDYFEFQLPGLKPNKPGS-----GDLVDAEGNVYGYTYTISEDGTVRTFNERI 314
Qy 105 NTKDDVKATLTMPAYIDPENVKKTGNVTLATGIGSTTANKTVLVDYKGYKFNLSIKGT 164
Db 315 TSESDIHGDFSLDTHLNDSDGRGPGDWVI--DIPTQEDLPVVIPI-----VPDT 362
Qy 165 IDQIDKTNNTYRQTIYVNPSPGDNVIAPVLTGNLKPNTDSNALIDQON----- 211
Db 363 EQQIDKQGHFDRTP--NPS-----AITWTV DIN--QAMKDQTNPTVTETWPTGNTF 409
Qy 212 TSIKVKVDNAAD--LSE-SYEVNPNFE-DVTNSVNIITFPNPNOYKVEFNTPPDDQITTP 267
Db 410 KSVKYVELVMNLNLDGTIKEVGRLESPDEYTVDKNGNVTIKGD TNKAYRLEYQTTIDEAVIP 469
Qy 268 ---YIVVVNGHI----DPNSKGDALRSTLYGYSNIIWRSMSWDN-----EVAFNNG 313
470 DGGDVPFKHNATLTSONPNGLDAEATVTATYGMKLDKRNIDYDEANOEFTWEINNYG 529

RESULT 13

US-09-002-285-94

; Sequence 94, Application US/09002285

; Patent No. 6369213

; GENERAL INFORMATION:

; APPLICANT: Schnepf, H. Ernest

; APPLICANT: Wicker, Carol

; APPLICANT: Narva, Kenneth E.

; APPLICANT: Walz, Michelle

; APPLICANT: Stockhoff, Brian

; APPLICANT: Muller-Cohn, Judy

; TITLE OF INVENTION: Toxins Active Against Pests

; NUMBER OF SEQUENCES: 105

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Saliwanchik, Lloyd & Saliwanchik

; STREET: 2421 N.W. 41st Street, Suite A-1

; CITY: Gainesville

; STATE: Florida

; COUNTRY: USA

; ZIP: 32606

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/002,285

; FILING DATE:

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/886,615

; FILING DATE: 1-JUL-1997

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/674,002

; FILING DATE: 1-JUL-1996

; ATTORNEY/AGENT INFORMATION:

; NAME: Sanders, Jay M.

; REGISTRATION NUMBER: 39,355

; REFERENCE/DOCKET NUMBER: MA-701C2

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (352) 375-8100

; TELEFAX: (352) 372-5800

; INFORMATION FOR SEQ ID NO: 94:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 789 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-09-002-285-94

Query Match 6.8%; Score 118; DB 4; Length 789;
Best Local Similarity 22.0%; Pred. No. 0.061;
Matches 82; Conservative 56; Mismatches 133; Indels 102; Gaps 18;
Qy 12 DITNQLTNVTYVGDSTTVYPHQAGYVKLNYSFVPSNAVSGDTFKITVPKELNLNGVTS 71
Db 148 EISKLDIINVNLINSTLTETPAYQRIKY---VNEKFELTFATETSSKVKKDG--S 201
Qy 72 TAKVPPIMAGDQVLANGVIDSGNVIYTFDYVNTKDDV-----KATLTMPAYIDP 122
Db 202 PADILDELAELTELAKSVTKND--VDGFEFYLNTFHDVMVGNLFGRSALKTASELITK 258
Qy 123 ENVKKTGN-----VTLATGIGS-----TTANKTV---LV DY-----EKYGKFY 157
Db 259 ENVKTSGSEVGNVNFVLVLTALQAKAFLTLTTCRKLGLADIDYTSIMNEHLNKEKEEF 318
Qy 158 NLSIKGTIDQIDKTNNTYRQTIYVNPSPGDNVIAPVLTGNLKPNTDSNALI--DQONT SIK 215
Db 319 RVNLLPTL-----SNTFSNPNTAKVKGSDAKMIV-EAKP---GHALIGFEISNDSIT 368
Qy 216 VYKVDNAADLSESYFVN-----PENFEDVTSVNITFPNPNOY---KV 255
Db 369 VLKV-YEAKLKQNYQVQDKDLSSEVIYGDMDKLLCPDQSEQIYYTNNIVF--PNEYVITKI 425
Qy 256 EF---NTPDDQITTPYIVVNVNGHIDPNKSGDLALRSTLYGYSNIIWRSMSWDNEVAFN 311
Db 426 DFTKKMKTTRYEVNTANFYDSTGEIDLNNK-----KVESSEAEYR 465
Qy 312 NGSGSGDGIDKPV 324
Db 466 TLSANDDDGVYMP 478

RESULT 14

US-09-206-942-65

; Sequence 65, Application US/09206942

; Patent No. 6432669

; GENERAL INFORMATION:

; APPLICANT: Loosmore, Sheena M.

; APPLICANT: Yang, Yan-Ping

; APPLICANT: Klein, Michel H.

; TITLE OF INVENTION: Protective Recombinant Haemophilus Influenzae High

; TITLE OF INVENTION: Molecular Weight Proteins

; FILE REFERENCE: 1038-861 MIS:jb

; CURRENT APPLICATION NUMBER: US/09/206,942

; CURRENT FILING DATE: 1998-12-08

; EARLIER APPLICATION NUMBER: 09/167,568

; EARLIER FILING DATE: 1998-10-07

; NUMBER OF SEQ ID NOS: 95

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 65

; LENGTH: 1180

; TYPE: PRT

; ORGANISM: Haemophilus influenzae

US-09-206-942-65

Query Match

Best Local Similarity 22.3%; Score 117; DB 4; Length 1180;

Matches 87; Conservative 49; Mismatches 143; Indels 112; Gaps 20;
Qy 3 AADAPAAAGTDITNQLTNVTG--IDSGTTVYPHQAGYVKLNYSFVPSNAVSGDTFKITV 60
Db 29 SADTPKRNNTKTTLTNSTLEKILARGSFVNITANNEIRVNSDINI-----GGNSHLTL 82
Qy 61 PKELNLNGVTSTAKVPPIMAGDQVLANGVIDSGNVIYTFDYVNTKDDVKATLTMPAYI 120
Db 83 WSSKNKN-----SGVLINGNITSTAN--GNLTISSGWDVHIKNITLES-- 124
Qy 121 DPENV-KKTGNVTLATGIGST-TANKTVLVDYKGYKFNLSIKGT-----IDQIDKT 171
Db 125 GRNLITTKEGDVAFKEKGNLNTITGQGTITAGNNKGRFENVSLNGTGLLENLSRPQKN 184
Qy 172 N-----NTYRQTIYVNPSPGD-NVIAPVLTGN-----LKPNTDSNALIDQON 211

Db 185 NSLVTNYFNGTLNIGSVNISMIPPNATSNWYSRYKGRYWNITHLNASDSNF-----N 239
Qy 212 TSIKYVKVDNAADLSESYFVNPFEDVTN-----SVNITFPNP-----NQYK----- 254
Db 240 LTIDSSAEDGSAPLLSSYTLNGISFTTDTTFNVNKNKNAFNFIKAPIGTINQYNNLNLYAL 299
Qy 255 -----VEFNTPDQITTPYIVVNVNGHIDPNSKGDALR-----STLYGYNS 295
Db 300 FNGNISVSGGGNVTFRLNASSNQOTPGVIINSKHLNA-SKGS-SLRFETTGSTKVGFELI 357
Qy 296 NIIWRMSWDNEVAFNNGSGS-----GDGID 321
Db 358 N-----NDLTLNATGGNISLLQVEGID 379

RESULT 15
US-09-206-942-63
; Sequence 63, Application US/09206942
; Patent No. 6432669
; GENERAL INFORMATION:
; APPLICANT: Loosmore, Sheena M.
; APPLICANT: Yang, Yan-Ping
; APPLICANT: Klein, Michel H.
; TITLE OF INVENTION: Protective Recombinant Haemophilus Influenzae High
; TITLE OF INVENTION: Molecular Weight Proteins
; FILE REFERENCE: 1038-861 MIS:jb
; CURRENT APPLICATION NUMBER: US/09/206,942
; CURRENT FILING DATE: 1998-12-08
; EARLIER APPLICATION NUMBER: 09/167,568
; EARLIER FILING DATE: 1998-10-07
; NUMBER OF SEQ ID NOS: 95
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 63
; LENGTH: 1188
; TYPE: PRT
; ORGANISM: Haemophilus influenzae
US-09-206-942-63

Query Match 6.8%; Score 117; DB 4; Length 1188;
Best Local Similarity 22.3%; Pred. No. 0.14;
Matches 87; Conservative 49; Mismatches 143; Indels 112; Gaps 20;
Qy 3 AADAPAGTDITNLTNTVVG--IDSGTTVYPHQAGYVKLNYSFVSPNSAVKGDFTKITV 60
Db 37 SADTPKRNNTKTTLTNSTLEKILARGSFVNITANNEIRVNSDINI-----GGNSHLTL 90
Qy 61 PKELNLNGVTSTAKVPPIMAGDQVLANGVIDSGNVIYFTDYVNTKDDVKATLTMPAYI 120
Db 91 WSSKNKN-----SGVLINGNITSTAN--GNLTIYSSGWVDIHKNITLES-- 132
Qy 121 DPENV-KKTGNVTLATIGIST-TANKTVLVDYKYGKFNLSIKGT-----IDQIDKT 171
Db 133 GRNLITTKEGDVAFEKGNLTLITGQGTITAGNNKGRFENVSLNGTGTGLLFLNLSRPQKN 192
Qy 172 N-----NTYRQTIYVNPNGD-NVIAPVLTGN-----LKPNTDSNALIDQON 211
Db 193 NSLVTNYFNGTLNIGSVNISMIPPNATSNWYSRYKGRYWNITHLNASDSNF-----N 247
Qy 212 TSIKYVKVDNAADLSESYFVNPFEDVTN-----SVNITFPNP-----NQYK----- 254
Db 248 LTIDSSAEDGSAPLLSSYTLNGISFTTDTTFNVNKNKNAFNFIKAPIGTINQYNNLNLYAL 307
Qy 255 -----VEFNTPDQITTPYIVVNVNGHIDPNSKGDALR-----STLYGYNS 295
Db 308 FNGNISVSGGGNVTFRLNASSNQOTPGVIINSKHLNA-SKGS-SLRFETTGSTKVGFELI 365
Qy 296 NIIWRMSWDNEVAFNNGSGS-----GDGID 321
Db 366 N-----NDLTLNATGGNISLLQVEGID 387

RESULT 16

US-09-002-285-96
; Sequence 96, Application US/09002285
; Patent No. 6369213
; GENERAL INFORMATION:
; APPLICANT: Schnepf, H. Ernest
; APPLICANT: Wicker, Carol
; APPLICANT: Narva, Kenneth E.
; APPLICANT: Walz, Michelle
; APPLICANT: Stockhoff, Brian
; APPLICANT: Muller-Cohn, Judy
; TITLE OF INVENTION: Toxins Active Against Pests
; NUMBER OF SEQUENCES: 105
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Saliwanchik, Lloyd & Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: Florida
; COUNTRY: USA
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/002,285
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/886,615
; FILING DATE: 1-JUL-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/674,002
; FILING DATE: 1-JUL-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Sanders, Jay M.
; REGISTRATION NUMBER: 39,355
; REFERENCE/DOCKET NUMBER: MA-701C2
; TELEPHONE: (352) 375-8100
; TELEFAX: (352) 372-5800
; INFORMATION FOR SEQ ID NO: 96:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 789 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-002-285-96

Query Match 6.7%; Score 116; DB 4; Length 789;
Best Local Similarity 22.0%; Pred. No. 0.091;
Matches 82; Conservative 57; Mismatches 132; Indels 102; Gaps 19;
Qy 12 DITNQLTNVTVGIDSGTTVYPHQAGYVKLNYSFVSPNSAVKGDFTKITVPKELNLNGVTS 71
Db 148 EISDKLDIINVNLINSLTFTITPAYQRIKY----VNEKFEELTFATETSSKVKKDG--S 201
Qy 72 TAKVPPIMAGDQVLANGVIDSGNVIYFTDYVNTKDDV-----KATLTMPAYIDP 122
Db 202 PADILDELTELTELAKSVTKND--VDGFEFYLTTFHDVVMGNNLFGRSALKTASELITK 258
Qy 123 ENVKKTGN-----VTLATIGIS-----TTANKTV---LVDY-----EKYKGFY 157
Db 259 ENVKTSGEVGNVNFLLIVLTAQAKAFLTLTTCRKLGLADIDYTSIMNEHLNKEKEEF 318
Qy 158 NLSIKGTIDQIDKTNNTYRQTIYVNPNGDNVIAPVLTGNLKPNTDSNALI--DQONTSIK 215
Db 319 RVNILPTL-----SNTFSPNPYAKVKGSEDEAKMIV-EAKP---GHALIGFEISNDSIT 368
Qy 216 YKVDNAADLSESYFVN-----PENFEDVTNSVNITFPNPQY---KV 255
Db 369 VLKV-YEAKLKQNYQVDKDSLSEVIYGDMDKLLCPDQSEQIYTTNNIVF--PNEYVITKI 425


```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 790 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: 36a
; US-09-073-898-4

Query Match      6.6%; Score 115; DB 4; Length 790;
Best Local Similarity 21.7%; Pred. No. 0.11;
Matches 81; Conservative 56; Mismatches 134; Indels 102; Gaps 18;

QY 12 DITNQLTNVTGIDSGTIVYPHQAGYVKLNLYGFSVPNSAVKGDTEFKITVPKELNNGVTS 71
Db 148 EISDKLDIINVNLINSLTTEITPAYQRIKY-----VNEKFEELTFATETSSKVKKDG--S 201

QY 72 TAKVPPIMAGDQVLANGVIDSDGNVIYFTTDYVNTKDDV-----KATLTMPAYIDP 122
202 PANILDELTELAKSVTKND---VDGFEFYLNTHFDVMVGNNLFGRSALKTASELITK 258

QY 123 ENVKKTGN-----VTLATGIGS-----TTANKTV---LVDY-----EKYGKFY 157
Db 259 ENVKTSGEVGNVYNFLIVLTALQAKAFLTLTTCRKLGLADIDYTSIMNEHLNKEKEEF 318

QY 158 NLSIKGTIDQIDKTNNTYRQTIYVNPSPGDNVIAPIVLTGNLKPNTDSNALI--DQONTSIK 215
Db 319 RVNLPPTL-----SNTFSNPYAKVKGSDAKMIV-EAKP---GHALIGFEISNDSIT 368

QY 216 VYKVDNAADLSESYFVN-----PENFEDVTNSVNITFPNPQY---KV 255
Db 369 VLKV-YEAKLKQNYQVDKSLSEVIYGMKLLCPDQSEQIYYTNNIVF--PNEYVITKI 425

QY 256 EF-----NTPDDQIITTPYIVVNVNGHIDPNKSGDLALRSTLYGYSNIIWRSMWDNEVAFN 311
Db 426 DFTKKMKTLRYEVTANFYDSSTGEIDLNNK-----KVESSEAEYK 465

QY 312 NGSGSGDGIDKPV 324
Db 466 TLSANDDGVYMP 478
```

```

RESULT 19
US-08-471-033-29
; Sequence 29, Application US/08471033
; Patent No. 5770696
GENERAL INFORMATION:
APPLICANT: Warren, Gregory W
APPLICANT: Koziel, Michael G
APPLICANT: Mullins, Martha A
APPLICANT: Nye, Gordon J
APPLICANT: Carr, Brian
APPLICANT: Desai, Nalini M
APPLICANT: Kostichka, N. Kristy
APPLICANT: Duck, Nicholas B
APPLICANT: Estruch, Juan J
TITLE OF INVENTION: No. 5770696el Pesticidal Proteins and Strains
NUMBER OF SEQUENCES: 50
CORRESPONDENCE ADDRESS:
ADDRESS: CIBA-GEIGY Corporation
STREET: 7 Skyline Drive
CITY: Hawthorne
STATE: NY
COUNTRY: USA
ZIP: 10532
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30B
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/471,033
```

```

; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/314,594
; FILING DATE: 09-SEP-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/218,018
; FILING DATE: 23-MAR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/037,057
; FILING DATE: 25-MAR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Pace, Gary M.
; REGISTRATION NUMBER: P-40,403
; REFERENCE/DOCKET NUMBER: CGC 1695/CIP3/DIV7 - SQLv3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8582
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 29:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 789 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-471-033-29

Query Match      6.6%; Score 114; DB 1; Length 789;
Best Local Similarity 21.7%; Pred. No. 0.14;
Matches 81; Conservative 56; Mismatches 134; Indels 102; Gaps 18;

QY 12 DITNQLTNVTGIDSGTIVYPHQAGYVKLNLYGFSVPNSAVKGDTEFKITVPKELNNGVTS 71
Db 148 EISDKLDIINVNLINSLTTEITPAYQRIKY-----VNEKFEELTFATETSSKVKKDG--S 201

QY 72 TAKVPPIMAGDQVLANGVIDSDGNVIYFTTDYVNTKDDV-----KATLTMPAYIDP 122
Db 202 PADILDELTELAKSVTKND---VDGFEFYLNTHFDVMVGNNLFGRSALKTASELITK 258

QY 123 ENVKKTGN-----VTLATGIGS-----TTANKTV---LVDY-----EKYGKFY 157
Db 259 ENVKTSGEVGNVYNFLIVLTALQAKAFLTLTTCRKLGLADIDYTSIMNEHLNKEKEEF 318

QY 158 NLSIKGTIDQIDKTNNTYRQTIYVNPSPGDNVIAPIVLTGNLKPNTDSNALI--DQONTSIK 215
Db 319 RVNLPPTL-----SNTFSNPYAKVKGSDAKMIV-EAKP---GHALIGFEISNDSIT 368

QY 216 VYKVDNAADLSESYFVN-----PENFEDVTNSVNITFPNPQY---KV 255
Db 369 VLKV-YEAKLKQNYQVDKSLSEVIYGMKLLCPDQSEQIYYTNNIVF--PNEYVITKI 425

QY 256 EF-----NTPDDQIITTPYIVVNVNGHIDPNKSGDLALRSTLYGYSNIIWRSMWDNEVAFN 311
Db 426 DFTKKMKTLRYEVTANFYDSSTGEIDLNNK-----KVESSEAEYR 465

QY 312 NGSGSGDGIDKPV 324
Db 466 TLSANDDGVYMP 478
```

```

RESULT 20
US-08-471-044-29
; Sequence 29, Application US/08471044
; Patent No. 5840868
GENERAL INFORMATION:
APPLICANT: Warren, Gregory W
APPLICANT: Koziel, Michael G
APPLICANT: Mullins, Martha A
APPLICANT: Nye, Gordon J
APPLICANT: Carr, Brian
APPLICANT: Desai, Nalini M
APPLICANT: Kostichka, N. Kristy
APPLICANT: Duck, Nicholas B
APPLICANT: Estruch, Juan J
```


;; PRIORITY APPLICATION DATA:
;; APPLICATION NUMBER: US 08/463,483
;; FILING DATE: 05-JUN-1995
;; PRIORITY APPLICATION DATA:
;; APPLICATION NUMBER: US 08/314,594
;; FILING DATE: 09-SEP-1994
;; PRIORITY APPLICATION DATA:
;; APPLICATION NUMBER: US 08/218,018
;; FILING DATE: 23-MAR-1994
;; PRIORITY APPLICATION DATA:
;; APPLICATION NUMBER: US 08/037,057
;; FILING DATE: 25-MAR-1993
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Meigs, J. Timothy
;; REGISTRATION NUMBER: 38,241
;; REFERENCE/DOCKET NUMBER: CGC1695/CIP3/DIV4 - SOLv4
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 919-541-8587
;; TELEFAX: 919-541-8689
;; INFORMATION FOR SEQ ID NO: 29:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 789 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
US-08-470-566B-29

Query Match 6.6%; Score 114; DB 2; Length 789;
Best Local Similarity 21.7%; Pred. No. 0.14;
Matches 81; Conservative 56; Mismatches 134; Indels 102; Gaps 18;

Qy 12 DITNQLTNVTGIDSGTTVYPHQAGYVKLNYSFVPSAVKGDTFKITVPKELNLNGVTS 71
Db 148 EISDKLDIINVNLINSLTTEITPAYQRIKY---VNEKFEELTFATETSSKVKKDG--S 201
Qy 72 TAKVPPIMAGDQVLANGVIDSGNVIYFTDYVNTKDDV-----KATLTMPAYIDP 122
Db 202 PADILDELTELTELAKSVTKND---VDGFEFYLNTFHDVMVGNLFGRSALKTASELITK 258
Qy 123 ENVKKTGN-----VTLATGIGS-----TTANKTV--LVDY-----EKYGFY 157
Db 259 ENVKTSGEVGNVYNFLIVLTAQAQAFLLTTCTCRKLLGLADIDYTSIMNEHLNKEKEEF 318
Qy 158 NLSIKGTIDQIDKNTTYRTIYVNPSPGDNVIAPVLTGNLKPNTDSNALI--DQQNTSIK 215
Db 319 RVNLIPTL-----SNTFSNPYAKVKGSDAKMIV-EAKP---GHALIGFEISNDSIT 368
Qy 216 VYKVDNAADLSESYFVN-----PENFEDVTSNVTNITFPNPQY---KV 255
Db 369 VLKV-YEAKLKQNYQVDKDSLEVIYGMCKLLCPDQSEQIYYTNNIVF--PNEYVITKI 425
Qy 256 EF----NTPDDQITTPYIVVNVNGHIDPNSKGDALRSTLYGNSNIWRSMSWDNEVAFN 311
Db 426 DFTKMKMKTLYREVTANFYDSSTGEIDLNKK-----KVESSEAEYR 465
Qy 312 NGSGSGDGIDKPV 324
Db 466 TLSANDDGVYMP 478

RESULT 24
US-08-838-219B-2
; Sequence 2, Application US/08838219B
; Patent No. 5877012
; GENERAL INFORMATION:
; APPLICANT: Warren, Gregory W
; APPLICANT: Koziel, Michael G
; APPLICANT: Mullins, Martha A
; APPLICANT: Nye, Gordon J
; APPLICANT: Carr, Brian
; APPLICANT: Desai, Nalini M
; APPLICANT: Kostichka, N. Kristy
; APPLICANT: Duck, Nicholas B

;; APPLICANT: Estruch, Juan J
;; TITLE OF INVENTION: A No. 5877012el Class of Proteins for the
;; TITLE OF INVENTION: Control of Plant Pests
;; NUMBER OF SEQUENCES: 20
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: CIBA-GEIGY Corporation
;; STREET: 7 Skyline Drive
;; CITY: Hawthorne
;; STATE: NY
;; COUNTRY: USA
;; ZIP: 10532
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.30B
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/838,219B
;; FILING DATE:
;; CLASSIFICATION: 530
;; PRIORITY APPLICATION DATA:
;; APPLICATION NUMBER: US 08/463,483
;; FILING DATE: 06-JUN-1995
;; PRIORITY APPLICATION DATA:
;; APPLICATION NUMBER: US 08/314,594
;; FILING DATE: 09-SEP-1994
;; PRIORITY APPLICATION DATA:
;; APPLICATION NUMBER: US 08/218,018
;; FILING DATE: 23-MAR-1994
;; PRIORITY APPLICATION DATA:
;; APPLICATION NUMBER: US 08/037,057
;; FILING DATE: 25-MAR-1993
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Pace, Gary M.
;; REGISTRATION NUMBER: 40,403
;; REFERENCE/DOCKET NUMBER: CGC 1925
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 919-541-8582
;; TELEFAX: 919-541-8689
;; INFORMATION FOR SEQ ID NO: 2:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 789 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
US-08-838-219B-2

Query Match 6.6%; Score 114; DB 2; Length 789;
Best Local Similarity 21.7%; Pred. No. 0.14;
Matches 81; Conservative 56; Mismatches 134; Indels 102; Gaps 18;

Qy 12 DITNQLTNVTGIDSGTTVYPHQAGYVKLNYSFVPSAVKGDTFKITVPKELNLNGVTS 71
Db 148 EISDKLDIINVNLINSLTTEITPAYQRIKY---VNEKFEELTFATETSSKVKKDG--S 201
Qy 72 TAKVPPIMAGDQVLANGVIDSGNVIYFTDYVNTKDDV-----KATLTMPAYIDP 122
Db 202 PADILDELTELTELAKSVTKND---VDGFEFYLNTFHDVMVGNLFGRSALKTASELITK 258
Qy 123 ENVKKTGN-----VTLATGIGS-----TTANKTV--LVDY-----EKYGFY 157
Db 259 ENVKTSGEVGNVYNFLIVLTAQAQAFLLTTCTCRKLLGLADIDYTSIMNEHLNKEKEEF 318
Qy 158 NLSIKGTIDQIDKNTTYRTIYVNPSPGDNVIAPVLTGNLKPNTDSNALI--DQQNTSIK 215
Db 319 RVNLIPTL-----SNTFSNPYAKVKGSDAKMIV-EAKP---GHALIGFEISNDSIT 368
Qy 216 VYKVDNAADLSESYFVN-----PENFEDVTSNVTNITFPNPQY---KV 255
Db 369 VLKV-YEAKLKQNYQVDKDSLEVIYGMCKLLCPDQSEQIYYTNNIVF--PNEYVITKI 425
Qy 256 EF----NTPDDQITTPYIVVNVNGHIDPNSKGDALRSTLYGNSNIWRSMSWDNEVAFN 311

Db 426 DFTKKMKTLYREVTANFYDSSTGEIDLNKK-----KVESSEAEYR 465

Qy 312 NGS GSGDGIDKPV 324

Db 466 TLSANDDGVY MPL 478

RESULT 25

US-08-469-334-29

; Sequence 29, Application US/08469334

; Patent No. 5990383

; GENERAL INFORMATION:

; APPLICANT: Warren, Gregory W

; APPLICANT: Koziel, Michael G

; APPLICANT: Mullins, Martha A

; APPLICANT: Nye, Gordon J

; APPLICANT: Carr, Brian

; APPLICANT: Desai, Nalini M

; APPLICANT: Kostichka, N. Kristy

; APPLICANT: Duck, Nicholas B

; APPLICANT: Estruch, Juan J

TITLE OF INVENTION: No. 5990383el Pesticidal Proteins and Strains

NUMBER OF SEQUENCES: 50

CORRESPONDENCE ADDRESS:

ADDRESSEE: CIBA-GEIGY Corporation

STREET: 7 Skyline Drive

CITY: Hawthorne

STATE: NY

COUNTRY: USA

ZIP: 10532

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30B

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/469,334

FILING DATE: 06-JUN-1995

CLASSIFICATION: 800

PRIOR APPLICATION NUMBER: 08/463,483

FILING DATE:

APPLICATION NUMBER: US 08/314,594

FILING DATE: 09-SEP-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/218,018

FILING DATE: 23-MAR-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/037,057

FILING DATE: 25-MAR-1993

ATTORNEY/AGENT INFORMATION:

NAME: Spruill, W. Murray

REGISTRATION NUMBER: 32,943

REFERENCE/DOCKET NUMBER: CGC 1695/CIP3

TELECOMMUNICATION INFORMATION:

TELEPHONE: 919-541-8615

TELEFAX: 919-541-8689

INFORMATION FOR SEQ ID NO: 29:

SEQUENCE CHARACTERISTICS:

LENGTH: 789 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-469-334-29

Query Match 6.6%; Score 114; DB 2; Length 789;

Best Local Similarity 21.7%; Pred. No. 0.14;

Matches 81; Conservative 56; Mismatches 134; Indels 102; Gaps 18;

Qy 12 DITNLTNVTVGIDSGTTVPYHQAGYVKLNFGFVSPNSAVKGDTEFKITVPKELNLNGVTS 71

Db 148 EISDKLDIINVNLINSLTETPAYQRIKY----VNEKFEELTFATETSSKVKKDG--S 201

Qy 72 TAKVPPIMAGDQVLANGVIDSGNVIYFTDYVNTKDDV-----KATLTMPAYIDP 122

Db 202 PADILDELTELAKSVTKND---VDGFEFYLNTFHDVMVGNLFGRSALKTASELITK 258

Qy 123 ENVKKTGN-----VTLATGIGS-----TTANKTV---LV DY-----EKY GKFY 157

Db 259 ENVKTS GSEVGNVNF LIVLTA LQAQAF LTLTTCRKLLGLADIDYTSIMNEHLNKEKEEF 318

Qy 158 NLSIKGTIDQIDKTNNTRYQRTIYVNPSPGDNVIA PVL TGNLKPNTDSNALI--DQONT SIK 215

Db 319 RVN ILPTL-----SNTFSNP NYAKVKGSD EDAKMIV-EAKP---GHALIGFEISNDSIT 368

Qy 216 VYKVDNAADLSESYFVN-----PENFEDVTNSVNI TFPNPQY---KV 255

Db 369 VLKV-YEAKLKQNYQVDKDSLSEVIYGDMDKLLCPDQSEQIYYTNNIVF--PNEYVITKI 425

Qy 256 EF----NTPDDQITTPYIVVYNGHIDPN SKGDLALRSTLYGYSNIIWRSMSWDNEVAFN 311

Db 426 DFTKKMKTLYREVTANFYDSSTGEIDLNKK-----KVESSEAEYR 465

Qy 312 NGS GSGDGIDKPV 324

Db 466 TLSANDDGVY MPL 478

Search completed: June 23, 2003, 14:04:51

Job time : 22.367 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 23, 2003, 14:03:36 ; Search time 46.9504 Seconds
(without alignments)
762.858 Million cell updates/sec

Title: US-10-056-052A-4
Perfect score: 1732
Sequence: 1 MVAADAPAAAGTDITNLTNV.....NGSGSGDGIDKPVVPEQPDE 331

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 417779 seqs, 108205813 residues
1 number of hits satisfying chosen parameters: 417779

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database : Published_Applications_AA:*

1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*

2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*

3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*

4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*

5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*

6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*

7: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*

8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*

9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*

10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*

11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*

12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*

13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*

14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1732	100.0	331	9	US-10-056-052-4
2	1732	100.0	345	10	US-09-813-820-7
3	1727	99.7	520	9	US-10-056-052-2
4	1723	99.5	936	1	US-08-781-986A-5249
5	1518	87.6	1021	10	US-09-815-242-5471
6	1518	87.6	1021	10	US-09-815-242-12544
7	409	23.6	767	10	US-09-815-242-5899
8	409	23.6	767	10	US-09-815-242-13140
9	346	20.0	978	10	US-09-815-242-5456
10	346	20.0	1001	10	US-09-815-242-12686
11	339	19.6	1018	10	US-09-815-242-5797
12	339	19.6	1018	10	US-09-815-242-12838
13	331	19.1	1027	1	US-08-781-986A-5254
14	282.5	16.3	251	1	US-08-781-986A-5252
15	257.5	14.9	1349	10	US-09-815-242-5898
16	257.5	14.9	1349	10	US-09-815-242-13137
17	232.5	13.4	932	10	US-09-815-242-5578
18	232.5	13.4	932	10	US-09-815-242-12438
19	210.5	12.2	841	10	US-09-815-242-5779

20	210.5	12.2	841	10	US-09-815-242-12751	Sequence 12751, A
21	136	7.9	559	1	US-08-781-986A-5251	Sequence 5251, Ap
22	130	7.5	900	9	US-09-884-465A-333	Sequence 333, App
23	130	7.5	906	9	US-09-884-465A-369	Sequence 369, App
24	130	7.5	906	9	US-09-884-465A-371	Sequence 371, App
25	130	7.5	906	9	US-09-884-465A-373	Sequence 373, App
26	129.5	7.5	1391	9	US-10-080-505-11	Sequence 11, Appli
27	125.5	7.2	921	9	US-09-117-447-6	Sequence 6, Appli
28	124.5	7.2	569	9	US-09-884-465A-235	Sequence 235, App
29	124.5	7.2	569	9	US-09-884-465A-255	Sequence 255, App
30	124.5	7.2	569	9	US-09-884-465A-256	Sequence 256, App
31	124.5	7.2	569	9	US-09-884-465A-356	Sequence 356, App
32	124.5	7.2	621	9	US-09-884-465A-368	Sequence 368, App
33	124.5	7.2	627	9	US-09-884-465A-366	Sequence 366, App
34	124.5	7.2	627	9	US-09-884-465A-367	Sequence 367, App
35	124.5	7.2	633	9	US-09-884-465A-349	Sequence 349, App
36	124.5	7.2	633	9	US-09-884-465A-350	Sequence 350, App
37	124.5	7.2	633	9	US-09-884-465A-351	Sequence 351, App
38	124.5	7.2	633	9	US-09-884-465A-352	Sequence 352, App
39	124.5	7.2	633	9	US-09-884-465A-353	Sequence 353, App
40	124.5	7.2	633	9	US-09-884-465A-354	Sequence 354, App
41	124.5	7.2	633	9	US-09-884-465A-355	Sequence 355, App
42	124.5	7.2	633	9	US-09-884-465A-357	Sequence 357, App
43	124.5	7.2	633	9	US-09-884-465A-358	Sequence 358, App
44	124.5	7.2	633	9	US-09-884-465A-359	Sequence 359, App
45	124.5	7.2	633	9	US-09-884-465A-360	Sequence 360, App
46	124.5	7.2	633	9	US-09-884-465A-361	Sequence 361, App
47	124.5	7.2	633	9	US-09-884-465A-362	Sequence 362, App
48	124.5	7.2	633	9	US-09-884-465A-363	Sequence 363, App
49	124.5	7.2	633	9	US-09-884-465A-364	Sequence 364, App
50	124.5	7.2	633	9	US-09-884-465A-365	Sequence 365, App
51	124.5	7.2	840	9	US-09-884-465A-10	Sequence 10, Appli
52	124.5	7.2	888	9	US-09-884-465A-338	Sequence 338, App
53	124.5	7.2	889	9	US-09-884-465A-348	Sequence 348, App
54	124.5	7.2	894	9	US-09-884-465A-336	Sequence 336, App
55	124.5	7.2	894	9	US-09-884-465A-337	Sequence 337, App
56	124.5	7.2	894	9	US-09-884-465A-340	Sequence 340, App
57	124.5	7.2	894	9	US-09-884-465A-342	Sequence 342, App
58	124.5	7.2	895	9	US-09-884-465A-344	Sequence 344, App
59	124.5	7.2	895	9	US-09-884-465A-346	Sequence 346, App
60	124.5	7.2	895	9	US-09-884-465A-347	Sequence 347, App
61	124.5	7.2	900	9	US-09-884-465A-334	Sequence 334, App
62	124.5	7.2	900	9	US-09-884-465A-335	Sequence 335, App
63	124.5	7.2	900	9	US-09-884-465A-339	Sequence 339, App
64	124.5	7.2	900	9	US-09-884-465A-341	Sequence 341, App
65	124.5	7.2	901	9	US-09-884-465A-343	Sequence 343, App
66	124.5	7.2	901	9	US-09-884-465A-345	Sequence 345, App
67	124.5	7.2	906	9	US-09-884-465A-332	Sequence 332, App
68	124.5	7.2	906	9	US-09-884-465A-370	Sequence 370, App
69	124.5	7.2	906	9	US-09-884-465A-372	Sequence 372, App
70	124.5	7.2	913	9	US-09-884-465A-384	Sequence 384, App
71	124.5	7.2	999	9	US-09-884-465A-376	Sequence 376, App
72	124.5	7.2	999	9	US-09-884-465A-377	Sequence 377, App
73	124.5	7.2	1039	9	US-09-884-465A-6	Sequence 6, Appli
74	124.5	7.2	1126	9	US-09-884-465A-383	Sequence 383, App
75	124.5	7.2	1139	9	US-09-884-465A-380	Sequence 380, App
76	124.5	7.2	1152	9	US-09-884-465A-379	Sequence 379, App
77	124.5	7.2	1238	9	US-09-884-465A-381	Sequence 381, App
78	124.5	7.2	1365	9	US-09-884-465A-382	Sequence 382, App
79	124.5	7.2	1378	9	US-09-884-465A-378	Sequence 378, App
80	123.5	7.2	1391	9	US-10-080-505-15	Sequence 15, Appli
81	123.5	7.1	1411	9	US-10-080-505-17	Sequence 17, Appli
82	121.5	7.0	5795	10	US-09-815-242-12610	Sequence 12610, A
83	118	6.8	789	9	US-10-099-285-94	Sequence 94, Appli
84	116	6.7	789	9	US-10-099-285-96	Sequence 96, Appli
85	116	6.7	839	9	US-10-023-437-23	Sequence 23, Appli
86	115	6.6	790	10	US-09-850-351A-4	Sequence 4, Appli
87	114	6.6	789	9	US-10-099-285-78	Sequence 78, Appli
88	114	6.6	789	9	US-10-099-285-100	Sequence 100, App
89	114	6.6	790	10	US-09-850-351A-8	Sequence 8, Appli
90	113.5	6.6	1228	9	US-09-117-447-2	Sequence 2, Appli
91	113	6.5	392	9	US-10-040-949A-46	Sequence 46, Appli
92	113	6.5	540	1	US-08-781-986A-5225	Sequence 5225, Ap

93 113 6.5 540 1 US-08-781-986A-5242 Sequence 5242, Ap
94 113 6.5 789 9 US-10-099-285-80 Sequence 80, Appl
95 113 6.5 789 9 US-10-099-285-90 Sequence 90, Appl
96 113 6.5 789 10 US-09-850-351A-6 Sequence 6, Appli
97 113 6.5 1394 9 US-09-839-996-2 Sequence 2, Appli
98 113 6.5 1394 9 US-10-080-505-2 Sequence 2, Appli
99 113 6.5 2478 10 US-09-815-242-5816 Sequence 5816, Ap
100 113 6.5 2478 10 US-09-815-242-12967 Sequence 12967, A

ALIGNMENTS

RESULT 1
US-10-056-052-4
; Sequence 4, Application US/10056052
; Publication No. US20030099656A1
; GENERAL INFORMATION:
; APPLICANT: PATTI, Joseph M
; APPLICANT: HUTCHINS, Jeff T
; APPLICANT: DOMANSKI, Paul
; APPLICANT: PATEL, Pratiksha
; APPLICANT: HALL, Andrea
; TITLE OF INVENTION: MONOCLONAL ANTIBODIES TO THE CLFA PROTEIN
; FILE REFERENCE: P07069US04/BAS
; CURRENT APPLICATION NUMBER: US/10/056,052
; CURRENT FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 60/308,116
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 60/298,413
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: 60/274,611
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: 60/264,072
; PRIOR FILING DATE: 2001-01-26
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 331
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-10-056-052-4

Query Match 100.0%; Score 1732; DB 9; Length 331;
Best Local Similarity 100.0%; Pred. No. 7.2e-126;
Matches 331; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MVAADAPAAAGTDITNQLTNVTGIDSGTTVYPHQAGYVKLYGFSVPNSAVKGDFTKITV 60
|||||
1 MVAADAPAAAGTDITNQLTNVTGIDSGTTVYPHQAGYVKLYGFSVPNSAVKGDFTKITV 60

61 PKELNLNGVTSTAKVPPIMAGDQVLANGVIDSDGNVIYFTDYVNTKDDVKATLTMPAYI 120
|||||
61 PKELNLNGVTSTAKVPPIMAGDQVLANGVIDSDGNVIYFTDYVNTKDDVKATLTMPAYI 120

121 DPENVKKTGNVTLATGIGSTTANKTVLVDYEKGKFNLSIKGTIDQIDKTNNTYRQTIY 180
|||||
121 DPENVKKTGNVTLATGIGSTTANKTVLVDYEKGKFNLSIKGTIDQIDKTNNTYRQTIY 180

181 VNPISGDNVIAVLTGNLKPNTDSNALIDQNTSIKVKYKVDNAADLSESYFVNPFEDVT 240
|||||
181 VNPISGDNVIAVLTGNLKPNTDSNALIDQNTSIKVKYKVDNAADLSESYFVNPFEDVT 240

241 NSVNITFPNPNOYKVEFNTPDDQITTPYIVVVNGHIDPNSKGDALRSTLYGYNNSNIWR 300
|||||
241 NSVNITFPNPNOYKVEFNTPDDQITTPYIVVVNGHIDPNSKGDALRSTLYGYNNSNIWR 300

301 SMSWDNEVAFNNGSGSGDGIDKPVVPEQPDE 331
|||||
301 SMSWDNEVAFNNGSGSGDGIDKPVVPEQPDE 331

RESULT 2

US-09-813-820-7
; Sequence 7, Application US/09813820
; Patent No. US20020102262A1
; GENERAL INFORMATION:
; APPLICANT: Hook, Magnus
; Patti, Joseph M.
; House-Pompeo, Karen
; Sthanam, Narayana
; Symersky, Jindrich
; TITLE OF INVENTION: COLLAGEN BINDING PROTEIN COMPOSITIONS
; AND METHODS OF USE
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: U.S.
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/813,820
; FILING DATE: 22-Mar-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/856,253
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Kitchell, Barbara S.
; REGISTRATION NUMBER: 33,928
; REFERENCE/DOCKET NUMBER: TAMK:193
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (512) 418-3000
; TELEFAX: (512) 474-7577
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 345 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-09-813-820-7

Query Match 100.0%; Score 1732; DB 10; Length 345;
Best Local Similarity 100.0%; Pred. No. 7.6e-126;
Matches 331; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVAADAPAAAGTDITNQLTNVTGIDSGTTVYPHQAGYVKLYGFSVPNSAVKGDFTKITV 60
|||||
Db 13 MVAADAPAAAGTDITNQLTNVTGIDSGTTVYPHQAGYVKLYGFSVPNSAVKGDFTKITV 72

QY 61 PKELNLNGVTSTAKVPPIMAGDQVLANGVIDSDGNVIYFTDYVNTKDDVKATLTMPAYI 120
|||||
Db 73 PKELNLNGVTSTAKVPPIMAGDQVLANGVIDSDGNVIYFTDYVNTKDDVKATLTMPAYI 132

QY 121 DPENVKKTGNVTLATGIGSTTANKTVLVDYEKGKFNLSIKGTIDQIDKTNNTYRQTIY 180
|||||
Db 133 DPENVKKTGNVTLATGIGSTTANKTVLVDYEKGKFNLSIKGTIDQIDKTNNTYRQTIY 192

QY 181 VNPISGDNVIAVLTGNLKPNTDSNALIDQNTSIKVKYKVDNAADLSESYFVNPFEDVT 240
|||||
Db 193 VNPISGDNVIAVLTGNLKPNTDSNALIDQNTSIKVKYKVDNAADLSESYFVNPFEDVT 252

QY 241 NSVNITFPNPNOYKVEFNTPDDQITTPYIVVVNGHIDPNSKGDALRSTLYGYNNSNIWR 300
|||||
Db 253 NSVNITFPNPNOYKVEFNTPDDQITTPYIVVVNGHIDPNSKGDALRSTLYGYNNSNIWR 312

QY 301 SMSWDNEVAFNNGSGSGDGIDKPVVPEQPDE 331
|||||

Db 313 SMSWDNEVAFNNGSGGDGIDKPVVPEQPDE 343

RESULT 3

US-10-056-052-2
; Sequence 2, Application US/10056052
; Publication No. US20030099656A1
; GENERAL INFORMATION:
; APPLICANT: PATTI, Joseph M
; APPLICANT: HUTCHINS, Jeff T
; APPLICANT: DOMANSKI, Paul
; APPLICANT: PATEL, Pratiksha
; APPLICANT: HALL, Andrea
; TITLE OF INVENTION: MONOCLONAL ANTIBODIES TO THE CLFA PROTEIN
; FILE REFERENCE: P07069US04/BAS
; CURRENT APPLICATION NUMBER: US/10/056,052
; CURRENT FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 60/308,116
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 60/298,413
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: 60/274,611
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: 60/264,072
; PRIOR FILING DATE: 2001-01-26
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 520
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-10-056-052-2

Query Match 99.7%; Score 1727; DB 9; Length 520;
Best Local Similarity 100.0%; Pred. No. 3.2e-125;
Matches 330; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	2	VAADAPAAAGTDITNQLTNVTGIDSGTTVYPHQAGYVKLVGFSVPNSAVKGDTEKITVP	61
Db	182	VAADAPAAAGTDITNQLTNVTGIDSGTTVYPHQAGYVKLVGFSVPNSAVKGDTEKITVP	241
Qy	62	KELNLNGVTSTAKVPPIMAGDQVLANGVIDSDGNVIYFTFDYVNTKDDVKATLTMPAYID	121
Db	242	KELNLNGVTSTAKVPPIMAGDQVLANGVIDSDGNVIYFTFDYVNTKDDVKATLTMPAYID	301
Qy	122	PENVKKTGNVTLATGIGSTTANKTVLVDYEKYGKFYNLSIKGTIDQIDKTNNTYRQTIYV	181
Qy	302	PENVKKTGNVTLATGIGSTTANKTVLVDYEKYGKFYNLSIKGTIDQIDKTNNTYRQTIYV	361
Qy	182	NPSGDNVIAPVLTGNLKPNTDSNALIDQQNTSIKVKYKVDNAADLSESYFVNPENFEDVTN	241
Db	362	NPSGDNVIAPVLTGNLKPNTDSNALIDQQNTSIKVKYKVDNAADLSESYFVNPENFEDVTN	421
Qy	242	SVNITFPNPQYKVEFNTPDQITTPYIVVNGHIDPNKSGDLALRSTLYGYSNIIWRS	301
Db	422	SVNITFPNPQYKVEFNTPDQITTPYIVVNGHIDPNKSGDLALRSTLYGYSNIIWRS	481
Qy	302	MSWDNEVAFNNGSGGDGIDKPVVPEQPDE	331
Db	482	MSWDNEVAFNNGSGGDGIDKPVVPEQPDE	511

RESULT 4

US-08-781-986A-5249
; Sequence 5249, Application US/08781986A
; Publication No. US20030054436A1
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5255
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue

CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/781,986A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Benson, Bob
REGISTRATION NUMBER: 30,446
REFERENCE/DOCKET NUMBER: PB248PP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 5249:
SEQUENCE CHARACTERISTICS:
LENGTH: 936 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-781-986A-5249

Query Match 99.5%; Score 1723; DB 1; Length 936;
Best Local Similarity 99.7%; Pred. No. 1.4e-124;
Matches 329; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy	2	VAADAPAAAGTDITNQLTNVTGIDSGTTVYPHQAGYVKLVGFSVPNSAVKGDTEKITVP	61
Db	230	VAADAPVAGTDITNQLTNVTGIDSGTTVYPHQAGYVKLVGFSVPNSAVKGDTEKITVP	289
Qy	62	KELNLNGVTSTAKVPPIMAGDQVLANGVIDSDGNVIYFTFDYVNTKDDVKATLTMPAYID	121
Db	290	KELNLNGVTSTAKVPPIMAGDQVLANGVIDSDGNVIYFTFDYVNTKDDVKATLTMPAYID	349
Qy	122	PENVKKTGNVTLATGIGSTTANKTVLVDYEKYGKFYNLSIKGTIDQIDKTNNTYRQTIYV	181
Db	350	PENVKKTGNVTLATGIGSTTANKTVLVDYEKYGKFYNLSIKGTIDQIDKTNNTYRQTIYV	409
Qy	182	NPSGDNVIAPVLTGNLKPNTDSNALIDQQNTSIKVKYKVDNAADLSESYFVNPENFEDVTN	241
Db	410	NPSGDNVIAPVLTGNLKPNTDSNALIDQQNTSIKVKYKVDNAADLSESYFVNPENFEDVTN	469
Qy	242	SVNITFPNPQYKVEFNTPDQITTPYIVVNGHIDPNKSGDLALRSTLYGYSNIIWRS	301
Db	470	SVNITFPNPQYKVEFNTPDQITTPYIVVNGHIDPNKSGDLALRSTLYGYSNIIWRS	529
Qy	302	MSWDNEVAFNNGSGGDGIDKPVVPEQPDE	331
Db	530	MSWDNEVAFNNGSGGDGIDKPVVPEQPDE	559

RESULT 5

US-09-815-242-5471
; Sequence 5471, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.

```
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5471
; LENGTH: 1021
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-5471
```

```
Query Match      87.6%; Score 1518; DB 10; Length 1021;
Best Local Similarity 87.3%; Pred. No. 1e-108;
Matches 288; Conservative 17; Mismatches 25; Indels 0; Gaps 0;

QY      2 VAADAPAAAGTDITNQLTNVTGIDSGTTVYPHQAGYVVKLVGFSVPNSAVKGDTEKITYP 61
      |||||||
Db      221 VAADAPAAAGTDITNQLTDVKVTIDSGTTVYPHQAGYVVKLVGFSVPNSAVKGDTEKITYP 280

QY      62 KELNLNGVTSTAKVPPIMAGDQVLANGVIDSDGNVIYTFDYYVNTKDDVKATLTMPAYID 121
      |||||||
Db      281 KELNLNGVTSTAKVPPIMVGDOVLANGVIDSDGNVIYTFDYYVNTKENVTANITMPAYID 340

QY      122 PENVKKTGNVTLATGIGSTTANKTVLVDYKEYGKFYNLSIKGTIDQIDKTNNTYRQTIYV 181
      |||||
Db      341 PENVTKTGNVTLTTGIGTNTASKTVLIDYKEYGQFHNLSIKGTIDQIDKTNNTYRQTIYV 400

QY      182 NPSGDNVIAPVLTGNLKPNTDSNALIDQQNTSIKVYKVDNAADLSESYFVNPFEDVTN 241
      |||||||
Db      401 NPSGDNVVLPALTGNLIPNTKSNALIDAKNTDIKVYRVDNANDLSESYVNPSPDFEDVTN 460

QY      242 SVNITFPNPQYKVEFNTPDQDITTPYIVVNGHIDPNKSGDLALRSTLYGYNNSNIWRS 301
      | :|||
Db      461 QVRISFPNANQYKVEFPTDDDDQITTPYIVVNGHIDPASTGDLALRSTFYGYDSNFIWRS 520

QY      302 MSWDNEVAFNNGSGSGDGIDKPVVPEQPDE 331
      |||||||
Db      521 MSWDNEVAFNNGSGSGDGIDKPVVPEQPDE 550
```

```
RESULT 6
US-09-815-242-12544
; Sequence 12544, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
```

```
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12544
; LENGTH: 1021
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-12544
```

```
Query Match      87.6%; Score 1518; DB 10; Length 1021;
Best Local Similarity 87.3%; Pred. No. 1e-108;
Matches 288; Conservative 17; Mismatches 25; Indels 0; Gaps 0;

QY      2 VAADAPAAAGTDITNQLTNVTGIDSGTTVYPHQAGYVVKLVGFSVPNSAVKGDTEKITYP 61
      |||||||
Db      221 VAADAPAAAGTDITNQLTDVKVTIDSGTTVYPHQAGYVVKLVGFSVPNSAVKGDTEKITYP 280

QY      62 KELNLNGVTSTAKVPPIMAGDQVLANGVIDSDGNVIYTFDYYVNTKDDVKATLTMPAYID 121
      |||||||
Db      281 KELNLNGVTSTAKVPPIMVGDOVLANGVIDSDGNVIYTFDYYVNTKENVTANITMPAYID 340

QY      122 PENVKKTGNVTLATGIGSTTANKTVLVDYKEYGKFYNLSIKGTIDQIDKTNNTYRQTIYV 181
      |||||
Db      341 PENVTKTGNVTLTTGIGTNTASKTVLIDYKEYGQFHNLSIKGTIDQIDKTNNTYRQTIYV 400

QY      182 NPSGDNVIAPVLTGNLKPNTDSNALIDQQNTSIKVYKVDNAADLSESYFVNPFEDVTN 241
      |||||||
Db      401 NPSGDNVVLPALTGNLIPNTKSNALIDAKNTDIKVYRVDNANDLSESYVNPSPDFEDVTN 460

QY      242 SVNITFPNPQYKVEFNTPDQDITTPYIVVNGHIDPNKSGDLALRSTLYGYNNSNIWRS 301
      | :|||
Db      461 QVRISFPNANQYKVEFPTDDDDQITTPYIVVNGHIDPASTGDLALRSTFYGYDSNFIWRS 520

QY      302 MSWDNEVAFNNGSGSGDGIDKPVVPEQPDE 331
      |||||||
Db      521 MSWDNEVAFNNGSGSGDGIDKPVVPEQPDE 550
```

```
RESULT 7
US-09-815-242-5899
; Sequence 5899, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
```


Db 310 DIKNIGDIKP---NNGETIATAKHDTANNLLITYTFTDYVDRFNSVQMGINYSIYMDADT 366
QY 125 VKKTGN-VTLATGIGSTTANKTVLVDYKEYGKFPYNLSIKG----TIDQIDKTNTT--YRQ 177
Db 367 IPVSKNDVEFNVTIGNTTTKTANIQYDPDYVVEKNSIGSAFTETVSHVGNKENPGYYKQ 426
QY 178 TIYVNPSGDNVIAPVLGTGNLKPNTDSNAL--IDQONTSIKVKYKVDNAADLSESYFVNPN 235
Db 427 TIYVNPSENLSNAKLKVQAYHSSYPNNIGQINKDVTDIKIYQVPKGYTLNKGYDVNTKE 486
QY 236 FEDVTSV--NITFPNPQYKVEFNTPDQITTPYIVVVGHIID-PNSKGDIALRSTLYG 292
Db 487 LTDVTNQYLQKITYGDNNSAVIDFGNAD----SAYVVMVNTKFQYTNSESPTLVQMATLS 542
QY 293 YNSNIIWRMSWDNEVAFNNGSGSGDG 319
Db 543 STGN--KSVSTGNALGFTNNQSGGAG 566

RESULT 16
US-09-815-242-13137
; Sequence 13137, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13137
; LENGTH: 1349
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-13137

Query Match 14.9%; Score 257.5; DB 10; Length 1349;
Best Local Similarity 26.9%; Pred. No. 2.1e-11;
Matches 88; Conservative 51; Mismatches 157; Indels 31; Gaps 12;
QY 13 ITNQLTNVTVGIDSGTTVYPHQAGYVKLNLYGFSVPNSAVKGDFTFKITVPKELNLNGVT-- 70
Db 251 ITSNTTLTVVDADKNNKIVPAQ-DYLSLKSQITVDDKVKSGDYFTIKYSDTVQVYGLNPE 309
QY 71 -----STAKVPPIMAGDQVLANGVIDSDGNVI-YTFTDYVNTKDDVKATLTMPAYIDPEN 124
Db 310 DIKNIGDIKP---NNGETIATAKHDTANNLLITYTFTDYVDRFNSVQMGINYSIYMDADT 366
QY 125 VKKTGN-VTLATGIGSTTANKTVLVDYKEYGKFPYNLSIKG----TIDQIDKTNTT--YRQ 177

Db 367 IPVSKNDVEFNVTIGNTTTKTANIQYDPDYVVEKNSIGSAFTETVSHVGNKENPGYYKQ 426
QY 178 TIYVNPSGDNVIAPVLGTGNLKPNTDSNAL--IDQONTSIKVKYKVDNAADLSESYFVNPN 235
Db 427 TIYVNPSENLSNAKLKVQAYHSSYPNNIGQINKDVTDIKIYQVPKGYTLNKGYDVNTKE 486
QY 236 FEDVTSV--NITFPNPQYKVEFNTPDQITTPYIVVVGHIID-PNSKGDIALRSTLYG 292
Db 487 LTDVTNQYLQKITYGDNNSAVIDFGNAD----SAYVVMVNTKFQYTNSESPTLVQMATLS 542
QY 293 YNSNIIWRMSWDNEVAFNNGSGSGDG 319
Db 543 STGN--KSVSTGNALGFTNNQSGGAG 566

RESULT 17
US-09-815-242-5578
; Sequence 5578, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5578
; LENGTH: 932
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-5578

Query Match 13.4%; Score 232.5; DB 10; Length 932;
Best Local Similarity 24.9%; Pred. No. 1.1e-09;
Matches 84; Conservative 55; Mismatches 158; Indels 41; Gaps 12;
QY 6 APAAGTDITNQL--TNVTVGIDSG-----TTVYPHQAGYVKLNLYGFSVPNSAVKGD 55
Db 169 APQOGTNVNDKVHFTNIDIAIDKGVHVKTTGNTFEWATSSDVLKLNKANYTTIDDSVKEGDT 228
QY 56 FKITVPKELNLNGVTSTAKVPPIM-AGDQVLANGVIDSD-GNVIYFTDYVNTKDDVKAT 113
Db 229 FTFKYQYFRPGSVRLPSQTQNLNAQGNIIAKGIYDSETSTTTTFTTNYVDQYTNISGS 288
QY 114 LTMPAYIDPEN--VKKTG---NVTLATGIGSTTANKTVLVDYKEYGKFPYNLSIKGTIDQI 168
Db 289 FEQVAFAKRENATTDKTAYPEVTL---GNDKYSKNVIVDYGKQKQQLTSSTNYINNE 344
QY 169 DKTNNTYRQTIYVNPSCDNIAPVLGTGNL---KPNTDSNALIDQQNTSIKVKYKVDNAADL 225

Db 345 DLSRN---MTVYVNPQPKKTYTKETFVNTLTGYKENPDAK-----NFKIYEVTNQNF 393

QY 226 SESYFVNPNFEDVTSVNIITFPNPQ-YKVEFNTPDDQITTPYIVVVNGHIDPNKSGDL 284

Db 394 VDSFTPDTSKLTVDTKFKITYSNDNKTATVDLLNGQSSDKQYIIQQVAYPDNSSTDNG 453

QY 285 ALRSTLYGYNNSIWRSMWDNEVAFNNGSGSGDGIDK 322

Db 454 KIDYTLETQNG-----KSSWSNSYSNVNGSSTANGDQK 486

RESULT 18

US-09-815-242-12438

; Sequence 12438, Application US/09815242

; Patent No. US20020061569A1

; GENERAL INFORMATION:

; APPLICANT: Haselbeck, Robert

; APPLICANT: Ohlsen, Kari L.

; APPLICANT: Zyskind, Judith W.

; APPLICANT: Wall, Daniel

; APPLICANT: Trawick, John D.

; APPLICANT: Carr, Grant J.

; APPLICANT: Yamamoto, Robert T.

; APPLICANT: Xu, H. Howard

; TITLE OF INVENTION: Identification of Essential Genes in

; TITLE OF INVENTION: Prokaryotes

; FILE REFERENCE: ELITRA.011A

; CURRENT APPLICATION NUMBER: US/09/815,242

; CURRENT FILING DATE: 2001-03-21

; PRIOR APPLICATION NUMBER: 60/191,078

; PRIOR FILING DATE: 2000-03-21

; PRIOR APPLICATION NUMBER: 60/206,848

; PRIOR FILING DATE: 2000-05-23

; PRIOR APPLICATION NUMBER: 60/207,727

; PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: 60/242,578

; PRIOR FILING DATE: 2000-10-23

; PRIOR APPLICATION NUMBER: 60/253,625

; PRIOR FILING DATE: 2000-11-27

; PRIOR APPLICATION NUMBER: 60/257,931

; PRIOR FILING DATE: 2000-12-22

; PRIOR APPLICATION NUMBER: 60/269,308

; PRIOR FILING DATE: 2001-02-16

; NUMBER OF SEQ ID NOS: 14110

; SOFTWARE: FastSEQ for Windows Version 4.0

; SEQ ID NO 12438

; LENGTH: 932

; TYPE: PRT

; ORGANISM: Staphylococcus aureus

US-09-815-242-12438

Query Match 13.4%; Score 232.5; DB 10; Length 932;

Best Local Similarity 24.9%; Pred. No. 1.1e-09;

Matches 84; Conservative 55; Mismatches 158; Indels 41; Gaps 12;

QY 6 APAAGTDITNQL--TNVTVGIDSG-----TTVYPHQAGYVVKLVGFSVPNSAVKGD 55

Db 169 APOQGTNVNDKVHFTNIDIAIDKGHVNKTGTEFWATSSDVLKLNKANYTIDDSVKEGDT 228

QY 56 FKITVPKELNLNGVTSTAKVPPIM-AGDQVLANGVIDSD-GNVIYFTDYVNTKDDVKAT 113

Db 229 FTFKYGYQYFRPGSVRLPSQTQNLNAOQNIAGIYDSETSTTTTYTFTNYVDQYTNISGS 288

QY 114 LTMPAYIDPEN--VKKTG---NVTLATGIGSTTANKTVLDYKEYGKFYNLSIKGTIDQI 168

Db 289 FEQVAFAKRENATFDKTAYPMEVTL----GNDKYSKNVIVDYGNGKQGQOLISSTNYINNE 344

QY 169 DKTNTYRQTIYVNPISGDNVIAPVLGTNL----KPNTDSNALIDQONTSIKVKVDNAADL 225

Db 345 DLSRN---MTVYVNPQPKKTYTKETFVNTLTGYKENPDAK-----NFKIYEVTNQNF 393

QY 226 SESYFVNPNFEDVTSVNIITFPNPQ-YKVEFNTPDDQITTPYIVVVNGHIDPNKSGDL 284

Db 394 VDSFTPDTSKLTVDTKFKITYSNDNKTATVDLLNGQSSDKQYIIQQVAYPDNSSTDNG 453

QY 285 ALRSTLYGYNNSIWRSMWDNEVAFNNGSGSGDGIDK 322

Db 454 KIDYTLETQNG-----KSSWSNSYSNVNGSSTANGDQK 486

RESULT 19

US-09-815-242-5779

; Sequence 5779, Application US/09815242

; Patent No. US20020061569A1

; GENERAL INFORMATION:

; APPLICANT: Haselbeck, Robert

; APPLICANT: Ohlsen, Kari L.

; APPLICANT: Zyskind, Judith W.

; APPLICANT: Wall, Daniel

; APPLICANT: Trawick, John D.

; APPLICANT: Carr, Grant J.

; APPLICANT: Yamamoto, Robert T.

; APPLICANT: Xu, H. Howard

; TITLE OF INVENTION: Identification of Essential Genes in

; TITLE OF INVENTION: Prokaryotes

; FILE REFERENCE: ELITRA.011A

; CURRENT APPLICATION NUMBER: US/09/815,242

; CURRENT FILING DATE: 2001-03-21

; PRIOR APPLICATION NUMBER: 60/191,078

; PRIOR FILING DATE: 2000-03-21

; PRIOR APPLICATION NUMBER: 60/206,848

; PRIOR FILING DATE: 2000-05-23

; PRIOR APPLICATION NUMBER: 60/207,727

; PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: 60/242,578

; PRIOR FILING DATE: 2000-10-23

; PRIOR APPLICATION NUMBER: 60/253,625

; PRIOR FILING DATE: 2000-11-27

; PRIOR APPLICATION NUMBER: 60/257,931

; PRIOR APPLICATION NUMBER: 60/269,308

; PRIOR FILING DATE: 2001-02-16

; NUMBER OF SEQ ID NOS: 14110

; SOFTWARE: FastSEQ for Windows Version 4.0

; SEQ ID NO 5779

; LENGTH: 841

; TYPE: PRT

; ORGANISM: Staphylococcus aureus

US-09-815-242-5779

Query Match 12.2%; Score 210.5; DB 10; Length 841;

Best Local Similarity 23.1%; Pred. No. 4.7e-08;

Matches 83; Conservative 57; Mismatches 136; Indels 83; Gaps 16;

QY 6 APAAGTDITNQL--TNVTVGIDSG-----TTVYPHQAGYVVKLVGFSVPNSAVKGD 55

Db 178 APOQGTNVNDKVHFSNIDIAIDKGHVNQTGKTEFWATSSDVLKLNKANYTIDDSVKEGDT 237

QY 56 FKITVPKELNLNGVTSTAKVPPIM-AGDQVLANGVIDSDGN-VIYFTDYVNTKDDVKAT 113

Db 238 FTFKYGYQYFRPGSVRLPSQTQNLNAOQNIAGIYDSTTTTYYTFTNYVDQYTNVRGS 297

QY 114 LTMPAYIDPEN--VKKTG---NVTLATGIGSTTANKTVLDYKEYGKFYNLSIKGTIDQI 168

Db 298 FEQVAFAKRKNATFDKTAYKMEVTL----GNDTYSEEIIVDYG-----NKKAQPLISST 347

QY 169 DKTNN---TYRQTIYVNPISGDNVIAPVLGTNL----KPNTDSNALIDQONTSIKVKVD 220

Db 348 NYINNEDLSRNMNTAYVNPQKNTYTKQTFVNTLTGYKFNPN-----KNFKIYEVT 397

QY 221 NAADLSESYFVNPNFEDVTSVNIITFPNP-----QYKV-EFNTPDDQ 263

Db 398 DQONQFVDSFTPDTSKLDKDVTDQDFDIYSNDNKTATVDLMKGQTSNNKQYIIQQVAYPDNS 457

QY 264 ITTPYIVVVNGHIDPNKSGDLALRSTLYGYNLSIWRSMWDNEVAFNNGSGSGDGIDK 322

Db 458 STD-----NGKIDYTLDTD-----KTKY-----SWSNSYSNVNGSSSTANGDQK 495

RESULT 20

US-09-815-242-12751

; Sequence 12751, Application US/09815242

; Patent No. US20020061569A1

; GENERAL INFORMATION:

; APPLICANT: Haselbeck, Robert

; APPLICANT: Ohlsen, Kari L.

; APPLICANT: Zyskind, Judith W.

; APPLICANT: Wall, Daniel

; APPLICANT: Trawick, John D.

; APPLICANT: Carr, Grant J.

; APPLICANT: Yamamoto, Robert T.

; APPLICANT: Xu, H. Howard

; TITLE OF INVENTION: Identification of Essential Genes in

; TITLE OF INVENTION: Prokaryotes

; FILE REFERENCE: ELITRA.011A

; CURRENT APPLICATION NUMBER: US/09/815,242

; PRIOR FILING DATE: 2001-03-21

; PRIOR APPLICATION NUMBER: 60/191,078

; PRIOR FILING DATE: 2000-03-21

; PRIOR APPLICATION NUMBER: 60/206,848

; PRIOR FILING DATE: 2000-05-23

; PRIOR APPLICATION NUMBER: 60/207,727

; PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: 60/242,578

; PRIOR FILING DATE: 2000-10-23

; PRIOR APPLICATION NUMBER: 60/253,625

; PRIOR FILING DATE: 2000-11-27

; PRIOR APPLICATION NUMBER: 60/257,931

; PRIOR FILING DATE: 2000-12-22

; PRIOR APPLICATION NUMBER: 60/269,308

; PRIOR FILING DATE: 2001-02-16

; NUMBER OF SEQ ID NOS: 14110

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 12751

; LENGTH: 841

; TYPE: PRT

; ORGANISM: Staphylococcus aureus

US-09-815-242-12751

Query Match 12.2%; Score 210.5; DB 10; Length 841;

Best Local Similarity 23.1%; Pred. No. 4.7e-08;

Matches 83; Conservative 57; Mismatches 136; Indels 83; Gaps 16;

6 APAAGTDITNQL--TNTVTGIDSG-----TTVYPHQAGYVVKLNYGFSVPNSAVKGD 55

178 APOQGTNVNDKVHFSNIDIAIDKGVHVNQTTGKTEFWATSSDVLLKLNKANYTIDDSVKEGDT 237

56 FKITVPKELNLNGVTSTAKVPPIM-AGDQVLANGVIDSDGN-VIYFTDYVNTKDDVKAT 113

238 FTFKYGOYFRPGSVRLPSQTQNLNNAQGNIIAKGIYDSTNTTYYTFTNYVDQYTNVRGS 297

114 LTMPIAYIDPEN--VKKTG---NVTLATGIGSTTANKTVLVVDYKEYKGFYNLSIKGTIDQI 168

298 FEQVAFAKRKNATTDKTAYKMEVTL----GNDTYSEEIIVDYG-----NKKAQPLISST 347

169 DKTNN--TYRQTIYVNPSPGDNVIAPVLGTNL-----KPNTDSNALIDQQNTSIKVKVD 220

348 NVINNEDLSRNMTAYVNPQPNKTYTKQTFVTNLTGKFNPNNA-----KNFKIYEVT 397

221 NAADLSESYFVNPFENFEDVTSVNITFPNPN-----QYKV-EFNTPPDDQ 263

398 DQNQFVDSFTPDTSKLKDVTQDFDVIYSNDNKTATVLMKQGTSSNKKYIIQQVAYPDNS 457

264 ITTPYIVVVNGHIDPNKSGDLALRSTLYGYNSNIWRMSWDNEVAFNNGSGSGDGIDK 322

458 STD-----NGKIDYTLDTD-----KTKY-----SWSNSYSNVNGSSSTANGDQK 495

US-08-781-986A-5251

; Sequence 5251, Application US/08781986A

; Publication No. US20030054436A1

; GENERAL INFORMATION:

; APPLICANT: Charles Kunsch

; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences

; NUMBER OF SEQUENCES: 5255

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Human Genome Sciences, Inc.

; STREET: 9410 Key West Avenue

; CITY: Rockville

; STATE: Maryland

; COUNTRY: USA

; ZIP: 20850

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage

; COMPUTER: HP Vectra 486/33

; OPERATING SYSTEM: MSDOS version 6.2

; SOFTWARE: ASCII Text

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/781,986A

; FILING DATE:

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER:

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

; NAME: Benson, Bob

; REGISTRATION NUMBER: 30,446

; REFERENCE/DOCKET NUMBER: PB248PP

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (301) 309-8504

; TELEFAX: (301) 309-8512

; INFORMATION FOR SEQ ID NO: 5251:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 559 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-781-986A-5251

Query Match 7.9%; Score 136; DB 1; Length 559;

Best Local Similarity 31.2%; Pred. No. 0.015;

Matches 29; Conservative 23; Mismatches 35; Indels 6; Gaps 3;

QY 236 FEDVTSNVNITFPNPNQYKVEFNTPDQITTPYIVVVNGHIDPNKSGDLALRSTLYGYNS 295

Db 14 FEDVTDNMSLDFDTNGGYSLNFNNLDQ--SKNYVIKVEGYDSDNA--SNLEFQTHLECYN 70

QY 296 NIWRMSWDNEVAFNNGSGSGDGIDK---PVV 325

Db 71 YYYTSNLTWKNGVAFYSNNAQGDGKDKLKEPII 103

RESULT 22

US-09-884-465A-333

; Sequence 333, Application US/09884465A

; Publication No. US20030077293A1

; GENERAL INFORMATION:

; APPLICANT: Shire Biochem, Inc.

; APPLICANT: Brodeur, Bernard

; APPLICANT: Martin, Denis

; APPLICANT: Charland, Nathalie

; APPLICANT: Ouellet, Catherine

; TITLE OF INVENTION: Streptococcus Antigens

; FILE REFERENCE: 055190-0044

; CURRENT APPLICATION NUMBER: US/09/884,465A

; CURRENT FILING DATE: 2001-06-20

; PRIOR APPLICATION NUMBER: 60/212,683

; PRIOR FILING DATE: 2000-06-20

; NUMBER OF SEQ ID NOS: 384

Db 450 ---SNSTLEEVPDPVQEKVAKFAESYGMKLENVLFNMDGTIELYL-----PSGEVIK 500
QY 189 -----IAPVLTGNLKPNTD---SNALIDQQNTSIKVKYKVDNAADLSESYFVNPNFE 237
Db 501 KNMAFTGEAPQNGENKPSGKNGKVGSTGVENQPTENK--PADSLPEAPNEKPVKPE--- 555
QY 238 DVTNSVNIFFPNPNQYKVEFNTPDQITTPYIVVVNGHIDP-----NSKGDALALRST 289
Db 556 ---NSTDNGMLNP-----EGNVGSDPMLDPALEEAPA-VDPVQEKLEKFTASYGLGLDSV 606
QY 290 LYGYNSNIIWRSMWDNEVAFNNGSGDGDIDKPVVPEQPD 331
Db 607 IFNMDGTIELRLPS--GEVIKKNLS---DFIAGPQITYTDDE 643

RESULT 25
US-09-884-465A-373
; Sequence 373, Application US/09884465A
; Application No. US20030077293A1
; GENERAL INFORMATION:
; APPLICANT: Shire Biochem, inc.
; APPLICANT: Hamel, Josee
; APPLICANT: Brodeur, Bernard
; APPLICANT: Martin, Denis
; APPLICANT: Charland, Nathalie
; APPLICANT: Ouellet, Catherine
; TITLE OF INVENTION: Streptococcus Antigens
; FILE REFERENCE: 055190-0044
; CURRENT APPLICATION NUMBER: US/09/884,465A
; PRIOR FILING DATE: 2001-06-20
; PRIOR APPLICATION NUMBER: 60/212,683
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 384
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 373
; LENGTH: 906
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Unknown Organism
US-09-884-465A-373

Query Match 7.5%; Score 130; DB 9; Length 906;
Best Local Similarity 23.6%; Pred. No. 0.083;
Matches 95; Conservative 47; Mismatches 152; Indels 108; Gaps 23;
1 MVAADAPAAAGTDITNQL-TNVTVGIDSGTIVYPHQAG--YVKLNYGFSVPNSA----- 50
279 IASKDYPEVSDGTFTVPTSLAYKMASQTIYPFHAGDTYLRVNPQFAVPKGTDALVRVF 338
QY 51 -----VKGDTFKITVPKELNNGVTSTA--KVPPIMAGDOVLANGVIDSDGNV 96
Db 339 DEFHGNAYLENNYKVGEIKLPIPK-LN-QGTRTAGNKIPVF-----MANAYLDNQSTY 391
QY 97 IYTF-----TD-----YVNTKDDVKATLTMPA-YIDPENVKKTGNVTLA 134
Db 392 IVEVPILEKENQTDKPSILPQFKRNKAQENSKLDEKVEEPTSEKVEKEKLSGTNST-- 449
QY 135 TGIGSTTANKTVLVD--YEKYGKF---YNLSIKGTIDQIDKTNNTYRQTIYVNPSPGDNV- 188
Db 450 ---SNSTLEEVPDPVQEKVAKFAESYGMKLENVLFNMDGTIELYL-----PSGEVIK 500
QY 189 -----IAPVLTGNLKPNTD---SNALIDQQNTSIKVKYKVDNAADLSESYFVNPNFE 237
Db 501 KNMAFTGEAPQNGENKPSGKNGKVGSTGVENQPTENK--PADSLPEAPNEKPVKPE--- 555
QY 238 DVTNSVNIFFPNPNQYKVEFNTPDQITTPYIVVVNGHIDP-----NSKGDALALRST 289
Db 556 ---NSTDNGMLNP-----EGNVGSDPMLDPALEEAPA-VDPVQEKLEKFTASYGLGLDSV 606
QY 290 LYGYNSNIIWRSMWDNEVAFNNGSGDGDIDKPVVPEQPD 331

Db 607 IFNMDGTIELRLPS--GEVIKKNLS---DFIAGPQITYTDDE 643
Search completed: June 23, 2003, 14:19:12
Job time : 48.9504 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

.. OM.protein - protein search, using sw model

Run on: June 23, 2003, 13:59:31 ; Search time 22.3014 Seconds
(without alignments)
1426.837 Million cell updates/sec

Title: US-10-056-052A-4
Perfect score: 1732
Sequence: 1 MVAADAPAAAGTDITNQLTNV.....NGSGSGDGIDKPVVPEQPDE 331

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues 283224

1 number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database : PIR_73:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES			
Result No.	Score	Query Match Length DB ID	Description
1	1727	99.7 933 2 S41539	fibrinogen-binding
2	1517	87.6 989 2 D89852	fibrinogen-binding
3	416.5	24.0 961 2 G90053	hypothetical prote
4	409	23.6 940 2 S19702	fibronectin-binding
5	349	20.2 1092 2 T30214	fibrinogen-binding
6	331	19.1 1018 2 A32192	fibronectin-binding
7	324.5	18.7 1038 2 H90053	hypothetical prote
8	323.5	18.5 1166 2 T28680	fibrinogen-binding
9	320.5	18.7 1141 2 E89824	hypothetical prote
10	312.5	18.0 877 2 F90070	Clumping factor B
11	257.5	14.9 1315 2 T28679	fibrinogen-binding
12	240.5	13.9 1385 2 D89824	hypothetical prote
13	229.5	13.3 953 2 C89824	hypothetical prote
14	182	10.5 1039 2 T30856	protein F2 - Strep
15	161.5	9.3 463 2 AG1542	wall associated pr
16	157.5	9.1 462 2 AH1184	wall associated pr
17	147.5	8.5 586 2 AD1458	probable peptidogl
18	141	8.1 1386 2 AC1533	surface protein (L
19	137	7.9 439 2 AE1251	probable peptidogl
20	137	7.9 4152 2 T31102	filamentous hemagg
21	135	7.8 1983 2 G86643	hypothetical prote
22	132.5	7.7 1457 2 D81019	adhesion and penet
23	132	7.6 1449 2 B81963	IgA-specific serin
24	129	7.4 762 2 A34355	cell surface prote
25	128	7.4 1158 2 AF1852	hypothetical prote
26	128	7.4 2021 2 A97859	190-KDa cell surfa
27	127	7.3 642 2 D81401	probable flagellar
28	127	7.3 793 2 AH1094	probable peptidogl
29	126.5	7.3 571 2 AI1094	probable peptidogl

30	126.5	7.3	1029	2	T30852	outer membrane pro
31	124.5	7.2	1039	2	D97985	hypothetical prote
32	123.5	7.1	1417	2	AG2137	hypothetical prote
33	122.5	7.1	657	2	AD1525	probable cell surf
34	122.5	7.1	1039	2	H95115	conserved hypothet
35	122.5	7.1	1487	2	AG2560	hypothetical prote
36	122	7.0	1578	2	AD1512	peptidoglycan boun
37	121.5	7.0	2249	2	A41477	190K surface antig
38	121.5	7.0	3890	2	C89921	hypothetical prote
39	121	7.0	1910	2	AF0394	probable adhesin h
40	119.5	6.9	691	2	B75622	hypothetical prote
41	119.5	6.9	836	2	D97182	extracellular neut
42	118.5	6.8	749	2	E86774	hypothetical prote
43	118.5	6.8	1369	2	T17504	hypothetical prote
44	118	6.8	1873	2	T30944	surface protein pr
45	117.5	6.8	449	2	JC7306	extracellular prot
46	116.5	6.7	725	2	E86790	hypothetical prote
47	116.5	6.7	1441	2	B86807	hypothetical prote
48	116.5	6.7	4199	2	S76412	hypothetical prote
49	116	6.7	784	2	AC1091	5'-nucleotidase, p
50	115.5	6.7	928	2	H86546	polymorphic outer
51	115.5	6.7	928	2	D72077	polymorphic outer
52	115.5	6.7	949	2	F81591	polymorphic membra
53	115.5	6.7	4936	2	AH2515	hypothetical prote
54	115	6.6	785	2	AB1455	5'-nucleotidase, p
55	115	6.6	1582	2	AC1153	adhesin homolog lm
56	114.5	6.6	1519	2	S41525	major ring-forming
57	114.5	6.6	1643	2	D71630	outer membrane pro
58	114.5	6.6	2269	2	T18472	hypothetical prote
59	114	6.6	1483	2	C97012	probably celluloso
60	114	6.6	1530	2	AH1396	peptidoglycan anch
61	114	6.6	2902	2	C71953	toxin-like outer m
62	114	6.6	3972	2	S75251	hypothetical prote
63	113.5	6.6	918	2	T02759	hypothetical prote
64	113.5	6.6	1228	2	I40468	surface layer prot
65	113.5	6.6	1250	2	G64993	yfaL protein - Esc
66	113.5	6.6	1902	1	B44858	lactocepin (EC 3.4
67	113.5	6.6	1902	2	B45764	lactocepin (EC 3.4
68	113.5	6.6	1962	2	A32634	lactocepin (EC 3.4
69	113	6.5	815	2	AB2444	hypothetical prote
70	113	6.5	1286	2	S28634	adhesin AIDA-I pre
71	113	6.5	1394	2	S60762	IgA-specific serin
72	112.5	6.5	634	2	F97172	flagellar hook-ass
73	112.5	6.5	1268	2	B99789	hemagglutinin/hemo
74	112.5	6.5	1270	2	E85649	hypothetical prote
75	112	6.5	649	2	S42894	metalloprotease
76	112	6.5	839	2	H97758	outer membrane ass
77	112	6.5	1902	2	S06997	lactocepin (EC 3.4
78	112	6.5	2481	2	D90011	FmtB protein [impo
79	111.5	6.4	980	2	H90681	probable flagellin
80	111.5	6.4	980	2	D85532	probable structura
81	111.5	6.4	1999	2	AB2018	hypothetical prote
82	111.5	6.4	3283	2	AC1018	large repetitive p
83	111	6.4	995	2	S50358	hypothetical prote
84	111	6.4	1036	2	T30311	S-layer protein -
85	111	6.4	4688	2	F82885	hypothetical prote
86	110.5	6.4	724	2	D96949	secreted protein c
87	110.5	6.4	859	2	AE2217	hypothetical prote
88	110.5	6.4	1176	2	A33856	surface-layer 125K
89	110.5	6.4	1566	2	T20058	hypothetical prote
90	110.5	6.4	1645	2	JN0896	crystalline surfac
91	110	6.4	410	2	AH1484	probable cell surf
92	110	6.4	1904	2	T13256	tail-host specific
93	110	6.4	5005	2	F82884	hypothetical prote
94	109.5	6.3	816	2	A71006	hypothetical prote
95	109	6.3	633	2	S61435	metalloprotease
96	109	6.3	1217	2	F97177	alpha-glucosidase
97	109	6.3	1579	2	B91290	probable invasiv
98	109	6.3	1700	2	G86131	probable invasiv
99	109	6.3	2044	2	AB1180	probable peptidogl
100	109	6.3	2340	2	B71704	cell surface antig

Db 447 THLSGYHKYPPYPYPVQLTWNNGVAFYSNNAKGDKKPNDP 493

RESULT 4

S19702

fibronectin-binding protein B - Staphylococcus aureus

C;Species: Staphylococcus aureus

C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 15-Oct-1999

C;Accession: S19702

R;Joensson, K.; Signaes, C.; Mueller, H.P.; Lindberg, M.

Eur. J. Biochem. 202, 1041-1048, 1991

A;Title: Two different genes encode fibronectin binding proteins in Staphylococcus aureus

A;Reference number: S19702; MUID:92111475; PMID:1837266

A;Accession: S19702

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-940 <JOE>

A;Cross-references: EMBL:X62992; NID:g49040; PIDN:CAA44726.1; PID:g581562

C;Keywords: fibronectin binding

Query Match 23.6%; Score 409; DB 2; Length 940;

Best Local Similarity 28.9%; Pred. No. 7.4e-18;

Matches 101; Conservative 72; Mismatches 131; Indels 46; Gaps 11;

QY 2 VAADAPAAAGTDITNQLTNVTVGIDSGT-----TVYPHQAGYVKLNLYGFSVPNSAV 51

Db 155 VVEETKATGTDVTK-----VEVEEGSEIVGHKQDTNVVNPNAERVTTLKYKVFEGGIK 209

QY 52 KGDTEKITVPKELNLNGVTSTAKVPPIMAGD-QVLANGVIDSDGNVIYTFDYVNTKDDV 110

Db 210 AGDYFDFTLSDNVETHGISTLRKVPEIKSTDGQVMATGEIIGERKVRYTEKEYVQEKD 269

QY 111 KATLTMPAYIDPENVKKTGNVTLATGIGSTTANKTVLVVDY----EKYKGKFNLSIKGTI 165

Db 270 TAELSLSNLFIDPTTVTQGNQNVKLVGETTVSKIFNIQYLGVRDNWG----VTANGRI 325

QY 166 DQIDKTNNTYRQTIYVNPSCDNIAPVLGTNL----KPNTDNALIDQQNTSIKVKVDN 221

Db 326 DTLNKKVDGKFSHFAYMKPNQSLSSVTVTGQVTKGNKPGVN-----NPTVKVYKHIG 377

QY 222 AADLSESYFVNPNEN---FEDVTNSVNITFPNPNQYKVEFTPDQITTPYIVVNVNGHIDP 278

Db 378 SDDLAEISVYAKLDDVSKFEDVTDNMSLDFDTNGGYSLNFNLDQ--SKNYVIKVEGYDYS 435

QY 279 NSKGDALRSTLYGYSNIIWRSMNSWDNEVAFNNGSGSGDGIDK---PVV 325

Db 436 NA-SNLEFQTHLFGYNYNYTSLNLTWKNGVAFYSNNAQGDGDKLKEPII 484

RESULT 5

T30214

fibrinogen-binding protein - Staphylococcus epidermidis

C;Species: Staphylococcus epidermidis

C;Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 11-May-2000

C;Accession: T30214

R;Nilsson, M.; Frykberg, L.; Flock, J.I.; Pei, L.; Lindberg, M.; Guss, B.

Infect. Immun. 66, 2666-2673, 1998

A;Title: A Fibrinogen-binding protein of Staphylococcus epidermidis.

A;Reference number: Z20781; MUID:98261511; PMID:9596732

A;Accession: T30214

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-1092 <NIL>

A;Cross-references: EMBL:Y17116; NID:e1296734; PID:e1296735; PIDN:CAA76638.1

Query Match 20.2%; Score 349; DB 2; Length 1092;

Best Local Similarity 30.3%; Pred. No. 5.1e-14;

Matches 100; Conservative 61; Mismatches 139; Indels 30; Gaps 14;

QY 13 ITNQLTNVTVGI-DSGTTVYPHQAGYVKLNLYGFSVPNSAVKGDTFKITVPKELNLNGVTS 71

Db 287 VTDQ--SITEGYDQSEGVIAKHAENLIYDVTFEVDKVKSGDVTMTVDIDKNTVPSDLTD 344

QY 72 TAKVPPIMAGD-QVLANGVIDS-DGNVIYTFDVTYNTKDDVKATLTMPAYIDPENVKKTG 129

Db 345 SFTIPKIKDNSGEIIATGTYDNKNKQITYTFTDYVDKYENIKAHLKLTYSIDKSKVPNNN 404

QY 130 ---NVTLATGIGSTTANKTVLVVDYKFKFYNLSTKGTIDQIDKTNNTYRQTIYVNPSCD 186

Db 405 TKLDVEYKTALSS--VNKTITVEYQRPNEERTANLQSMFTNIDTKNHTVEQTIYNPL-- 460

QY 187 NVIAPVLGTGNLKPNTDSNALIDQQNTSIKVKYKVDNAADLSESYFV-NPENFEDVTNSVNI 245

Db 461 RYSAKETNVNISNGSGESTIIDSTIIKVKYKVGDNQNLPSNRIDYSEYEDVTNDIYA 520

QY 246 TFPNPNQYKVEFNTPDQITTPYIVVNVNGHIDPNKSGD-----LALRSTLYGYSNII 298

Db 521 QLGNNNDVNINFG----NIDSPYIIKVISKYDPN-KDDYTTIQQTVTMTQTINEYTGE-- 573

QY 299 WRSMNSWDNEVAFNNGSGSGDGIDKPVVPEQ 328

Db 574 FRTASYDNTIAFSTSSGQGG-DLP--PEK 600

RESULT 6

A32192

fibronectin-binding protein - Staphylococcus aureus

C;Species: Staphylococcus aureus

C;Date: 31-Jul-1989 #sequence_revision 31-Jul-1989 #text_change 16-Feb-1997

C;Accession: A32192

R;Signaes, C.; Raucchi, G.; Joensson, K.; Lindgren, P.E.; Anantharamaiah, G.M.; Hoeoek

Proc. Natl. Acad. Sci. U.S.A. 86, 699-703, 1989

A;Title: Nucleotide sequence of the gene for a fibronectin-binding protein from Staph

A;Reference number: A32192; MUID:89098998; PMID:2521391

A;Accession: A32192

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-1018 <SIG>

A;Cross-references: GB:J04151

C;Keywords: fibronectin binding

Query Match 19.1%; Score 331; DB 2; Length 1018;

Best Local Similarity 24.6%; Pred. No. 6.1e-13;

Matches 82; Conservative 77; Mismatches 146; Indels 28; Gaps 11;

QY 10 GTDITNQLTNVTVGIDSG---TTVYPHQAGYVKLNLYGFSVPNSAVKGDTFKITVPKELN 65

Db 194 GTDVTSKVT-VEIGSIEGHNTNKVEPHAGQRAVLKYKLKFENGLHQGDYDFDTLSNNVN 252

QY 66 LMGVTSTAKVPPIMAGDQVLANGVIDSDGNVIYTFDYVNTKDDVKATLTMPAYIDPENV 125

Db 253 THGVSTARKVPEIKNGSVVMATGEVLEGKIRYTFETNDIEDKVDVTAELEINLFDPKTV 312

QY 126 KKTGNVTLATGIGSTTANKTVLVVDY-EKYGKFYNLSIKGTIDQIDKTNNTYRQTIYVNP 184

Db 313 QTNGNQITITSTLNEEQTSKELDVKYKDGIGNYY-ANLNGSIETFNKANNRFSHVAFKPN 371

QY 185 GDNVIAPVLGTGNLKPNTDSNALIDQQNTSIKVKYK-VDNAADLSESYFVN---PENFEDVT 240

Db 372 NGKTTTSVTVTGTLMKGSNQNG----NQPKVRIFEYLGNNEDIAKSVYANTTDTSKFKEVT 427

QY 241 NSV--NITFPNPNQYKVEFTPDQITTPYIVVNVNGHIDPNKSGDLALRSTLYGYSNII 298

Db 428 SNMCSNLNLQNGSYSLNI-----ENLDKTYVVVHYDGEY-LNGTDEVDERTQMVGHPDQLY 482

QY 299 -----WRSMNSWDNEVAFNNGSGSGDGIDKPVV 325

Db 483 KYYYDRGYTLTDWNGLVLYSNKANGNEKNGP 515

RESULT 7

H90053

hypothetical protein fnb [imported] - Staphylococcus aureus (strain N315)

C;Species: Staphylococcus aureus

C;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001

C;Accession: H90053

Db 234 KKVLSSTN--VQPIALDSNRNFELTIGTLN-----NQSVV--ITYDTKITTKQKSYTN-- 281

QY 276 IDPNSKGDALRSTLYGYNSSNIWRSMWDNEVAFNNGSGGDGIDKPVVPEQP 329

Db 282 -----KATLSGDNLDVAVSRNATVND--YSGGGQGTGTPPTPPVKEEP 322

RESULT 16

AH1184

wall associated protein precursor (LPXTG motif) homolog lmo0880 [imported] - Listeria monocytogenes

C;Species: Listeria monocytogenes

C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001

C;Accession: AH1184

R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker, P.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.; Jones, L.M.; Karst, U.

Science 294, 849-852, 2001

A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.; Jones, L.M.; Karst, U.

Science 294, 849-852, 2001

A;Title: Comparative genomics of Listeria species.

A;Reference number: AB1077; MUID:21537279; PMID:11679669

A;Accession: AH1184

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-462 <GLA>

A;Cross-references: GB:NC_003210; PIDN:CAC98958.1; PID:g16410283; GSPDB:GN00177

A;Experimental source: strain EGD-e

C;Genetics:

A;Gene: lmo0880

Query Match 9.1%; Score 157.5; DB 2; Length 462;

Best Local Similarity 22.0%; Pred. No. 0.014;

Matches 80; Conservative 52; Mismatches 123; Indels 109; Gaps 16;

QY 9 AGTDI-TNQLTNVTGIDSGTTVYPHQA-----GYVKLNYGFSVPNSAVKGDTEKITVP 61

Db 25 AATDYGSSFTTNSLQNG-----EQATNFKENSKVRVAYDFVITQPVASGETMTLTIP 79

QY 62 KELNLNGVTSTAKVPPIMAGDQVLANGVIDS-DGNVIYFTFDYVNTKDDVKATLTMPAYI 120

Db 80 DQLKL---INYGFPPLMDSQGTIANATIDQVGTITLTFTDYVNTHTDLGSLFYNATF 136

QY 121 DPNVKKTKGNVTLATGIGSTTANKTVLVD-----YEKYGK-----YN 158

Db 137 NSKNIQTDOVNPIAPVPKNTTQTVPYISKVNSGGGTGSPITVFKQGRMDKDLILHWT 196

QY 159 LSIKGTIDQIDKTNNTYRQTIYVNPSPGDNVIA-----PVLTKNLKPNWDSNALI 207

197 VTLNNAITPID--NAVYDTL--GSGQNLGSAIKYRDANKKVIATNIQP----IALD 247

QY 208 DQNTSIKVKYKVDNAADLSESYFVNPFEDVTNSVNIITFPNPNQYKVEFNPDDQITTP 267

Db 248 ADRNFELSIGNALN-----QSVVITY-----DTKITTK 275

QY 268 YIVVNGHIDPNKSGDLALRSTLYGYNSSNIWRSMWDNEVAFNNGSGGDGIDKPVVPE 327

Db 276 QKSYTN-----KATLSGDNLDVSR-----NATVNDYSGG-CQGTGTPPAPP 316

QY 328 QPDE 331

Db 317 VKEE 320

RESULT 17

AD1458

probable peptidoglycan bound protein (LPXTG motif) lin0203 [imported] - Listeria innocua

C;Species: Listeria innocua

C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001

C;Accession: AD1458

R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker, P.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.; Jones, L.M.; Karst, U.

Science 294, 849-852, 2001

Query Match 8.1%; Score 141; DB 2; Length 1386;

Best Local Similarity 20.1%; Pred. No. 0.69;

Matches 84; Conservative 40; Mismatches 117; Indels 176; Gaps 14;

A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehla

A;Title: Comparative genomics of Listeria species.

A;Reference number: AB1077; MUID:21537279; PMID:11679669

A;Accession: AD1458

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-586 <GLA>

A;Cross-references: GB:AL592022; PIDN:CAC95436.1; PID:g16412622; GSPDB:GN00178

A;Experimental source: strain Clip11262

C;Genetics:

A;Gene: lin0203

Query Match 8.5%; Score 147.5; DB 2; Length 586;

Best Local Similarity 25.4%; Pred. No. 0.084;

Matches 85; Conservative 36; Mismatches 117; Indels 97; Gaps 17;

QY 36 GY---VKLNYGFSVPNSA-VK-GDTFKITVPKELNLNGVTSTAKVPPIMAGDQVLANGVI 90

Db 62 GYDVSQVHYHTWALPNSNNVAGDTMQFVLPPELQI--VTDLDFSLKDHKGN-VVGNVVA 118

QY 91 DSD-GNVIYFTFDYVNTKDDVKATLTMPAYIDPNVKKTKGNVTLATGIGSTTANKTVLVD 149

Db 119 TKDTGKVVIITFDEKNSDVSGLDFWSNWDKSLVEGNEKVPVFPVNGTITVEV- 177

QY 150 YEKYGKFNLSIKGTIDQIDKTNNTYRQTIYVNPSPGDNVIAVLTGNLKNPNTDSNALIDQ 209

Db 178 -----GGKNQIDPTETLYKYG-WANAKNPELIQVVRVNYAKENIQNAVYED 223

QY 210 QNTSIKVKYKVDNAADLSESYFVNPFEDVTNSVNIITFPNPNQYKVEFNPDDQITTPYI 269

Db 224 -----FVGPKQVIDF-NSIQAVHG-----EFD-PDDNFTp--- 251

QY 270 VVNGHIDPNKSGDLALRSTLYGYNSSN-----IWRSMWDN----- 306

Db 252 ----GAAPSSD----IIQTTDGFKVNLGNTLDSVKISYYTTSTONGASPSYTNKGKLTG 303

QY 307 -----EVA--FNNGSGSGDGIDKPVVPEQDPE 331

Db 304 DNYVTQIEVATPTSGSGGGEGTGTGVELTKTDD 338

RESULT 18

AC1533

surface protein (LPXTG motif) [imported] - Listeria innocua (strain Clip11262)

C;Species: Listeria innocua

C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001

C;Accession: AC1533

R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloec

.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi,

D.; Jones, L.M.; Karst, U.

Science 294, 849-852, 2001

A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.;

Ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehla

A;Title: Comparative genomics of Listeria species.

A;Reference number: AB1077; MUID:21537279; PMID:11679669

A;Accession: AC1533

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-1386 <GLA>

A;Cross-references: GB:AL592022; PIDN:CAC96035.1; PID:g16413254; GSPDB:GN00178

A;Experimental source: strain Clip11262

C;Genetics:

A;Gene: lin0803

Query Match 8.1%; Score 141; DB 2; Length 1386;

Best Local Similarity 20.1%; Pred. No. 0.69;

Matches 84; Conservative 40; Mismatches 117; Indels 176; Gaps 14;

QY 29 TVYPHQAGYVKLNYGFS-VPNSAVKGDTEKITVPKELNLNGVTSTAKVPPIMAGDQVLAN 87

Db 924 TIDTTKAGNYDITYSGVTRSTELSKTITVTVK-----NQVNLEAKDSTLYEGDKWIAK 979

Qy	88	----GVIDSDGN-----	95
Db	980	DNFVSATDKDGNVTDFKAIEVKGTVNTTKAGTYKITYSYAGISKTTITVTVLANQTKIVAK	1039
Qy	96	--VIYFTDYVNTKDDVKATLTMTPAYIDPENVKKTGNVTLAT-----	GIGSTTA 142
Db	1040	DLTIYEGDNWKEQDNFVSATDKFGQAIDFNSVKVTGSDVIQTPGKYRITYSIEG-ASTTI	1098
Qy	143	NKTVLVDY-----	EKGKFEYNLS----IKGTIDQIDKTN 173
Db	1099	TVTVLADQSNLVAKNSTIYVGDKWQSKDNFVSATDKYGPIDLSLLTVTGIVD--TTTPG	1156
Qy	174	TYRQTIYVN-----	PSGDNVIAPV----- 192
Db	1157	EYEITYSVNGLTTTITVTVKENQANIVAEDSTIHAKESWKAADNFVSATDKTGKNIGLSS	1216
Qy	193	--LTGNLKPNTDSN-----	ALIDQQNTSIKV---YKVDNAADL 225
Db	1217	VSVTGKVDINTPGNYEITYITIDGVSTTITVTVLVNHQSIEAHDAKIKVGTSWKPEDSFIL	1276
Db	226	SESYFVNPFENFEDVTVNSVNIITFPNPNOYKVEFNTPDDQITTPYIVVNVNGHIDPNSKG	282
Db	1277	AKDKFGETADFSVTVGTVDTTKPGKYQIITYTIDGVSVTITVIV-----	EDNSKG 1327
RESULT 19			
AE1251			
probable peptidoglycan bound protein (LPXTG motif) lmo1413 [imported] - Listeria monocytogenes			
C;Species: Listeria monocytogenes			
C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001			
C;Accession: AE1251			
R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker, J.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.D.; Jones, L.M.; Karst, U.			
Science 294, 849-852, 2001			
A;Authors: KrefT, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Mak, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, A.; Title: Comparative genomics of Listeria species.			
A;Reference number: AB1077; MUID:21537279; PMID:11679669			
A;Accession: AE1251			
A;Status: preliminary			
A;Molecule type: DNA			
A;Residues: 1-439 <GLA>			
A;Cross-references: GB:NC_003210; PIDN:CAC99491.1; PID:g16410842; GSPDB:GN00177			
A;Experimental source: strain EGD-e			
C;Genetics:			
A;Gene: lmo1413			
Query Match 7.9%; Score 137; DB 2; Length 439;			
Best Local Similarity 23.5%; Pred. No. 0.25;			
Matches 81; Conservative 48; Mismatches 128; Indels 88; Gaps 20;			
Qy	31	YPHQAGYVKL-NYGFSPNSAVKGD-TFKITVPKEL-----NLNGVTSTAKV--PPIMAGD	82
Db	63	YPGNAAYVLFNNKQYNITDYTVENDVNFRLKLPKTLKLEAGDTLNYFTITGNVLDPVAYPGQ	122
Qy	83	Q--VLANGVIDSD-GNVIYFTDYVN---TKDDVKATL-----TMPAYIDPENVKKTGN	130
Db	123	ESYKLAGPFSPIEKANIQVNVVDETNTQTLATSDRLSGKLGETYTTTSPKAIDGYQVKET--	180
Qy	131	VTLATGIGSTTANKTVLVLDYEKYGFYNLSIKGTIDQIDKTNNTYRQTIYVNPSPGDNVIA	190
Db	181	PTNATGT-YTTNTETIQINY---VYEKTAVEGANVSVE-----YINEATNESIA	225
Qy	191	P--VLTGNL-----	KPNTDSNALIDQQNTSIKVYKVDNAADLSESYF 230
Db	226	PTETLSGKIGTTTFOAEVKEIDGYELSQVPSNQSGTYTDQTSQVIFKYKKITTPD-----	279
Qy	231	VNPFENFEDVTVNSVNIITFPNPNOYKVEFNTPDDQITTPYIVVNVNGHI----	DPNSKGD LAL 286
Db	280	-----EVAKDVTVT-----YK---DTKGNQLADP--VILKGDIGSTFETEAKSFKGY	321
Qy	287	RSTLYGYNSTIIWRSMSWDNEVAFNNGSGSGGDGIDKPVVPEQPD	331

Db	322	KLTKTPSNHSSVFTSDSQEVEYVYSKDDAV---ITPPVTPVNPDK	363
RESULT 20			
T31102			
filamentous hemagglutinin 1 - Haemophilus ducreyi			
C;Species: Haemophilus ducreyi			
C;Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 22-Oct-1999			
C;Accession: T31102			
R;Ward, C.K.; Lumbley, S.R.; Latimer, J.L.; Cope, L.D.; Hansen, E.J.			
J. Bacteriol. 180, 6013-6022, 1998			
A;Title: Haemophilus ducreyi secretes a filamentous hemagglutinin-like protein.			
A;Reference number: Z20984; MUID:99030326; PMID:9811662			
A;Accession: T31102			
A;Status: preliminary; translated from GB/EMBL/DBJ			
A;Molecule type: DNA			
A;Residues: 1-4152 <WAR>			
A;Cross-references: EMBL:AF057695; NID:g3929017; PID:g3929018; PIDN:AAC79757.1			
C;Genetics:			
A;Gene: lspA1			
Query Match 7.9%; Score 137; DB 2; Length 4152;			
Best Local Similarity 22.0%; Pred. No. 5.6;			
Matches 72; Conservative 55; Mismatches 129; Indels 72; Gaps 15;			
Qy	4	ADAPAAAGTDITNLTNVTVGIDS-----GTTVYPHQAGYVKLNYGFSPNSAVKGDTFKI	58
Db	619	SDFTSNGSKLVDAQNNLTNVNFNFNITQGSEIILH--GNVTLNAKGNTNSG-----NL	670
Qy	59	TVPKELNLNGVTSTAKVPPIMAGD--QVLANGVIDSDGNVIYFTDYVNTKDDVKATLTM	116
Db	671	TTMKELNISNIESFINAGNLTTGKNLEVHSNTTVKNDKGLVSIENLNISSTKDTFTNNGT	730
Qy	117	PAYIDPENVKKTGNVTLATGIGSTTANKTVLVDEYKGFYNLSIKGTIDQIDKTNNTYR	176
Db	731	LG-LEALKIASGGNFTNASN-GSLASNKSL----DIYGN--NFTNNGTIESVKSLNITNN	782
Qy	177	QTIYVNPSPGDNVIAPVLTGNLKPNTDSNALIDQQNTSIKVYKVDNAADLSESYFVNPN	235
Db	783	YT-FIN-----	NATIKSYGLNIT--SQGNFTNDSNG 811
Qy	236	-----FEDVTVNSVNITFPNP-----QYKVEFNTPDDQITTPYIVV-VNGHIDPNS	280
Db	812	TVMSHDLNLNITSQANII--NKNLLAGGQGLNLTAKNITNDSNSTAIAVLHNSNDINLNA	869
Qy	281	KGDLALRSTLYGYNSTIIWRSMSWDNEV	308
Db	870	NNKVYNIGEIYSQAGNISVEAKLLHNDV	897
RESULT 21			
G86643			
hypothetical protein ybEF [imported] - Lactococcus lactis subsp. lactis (strain IL140			
C;Species: Lactococcus lactis subsp. lactis			
C;Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 03-Aug-2001			
C;Accession: G86643			
R;Bolotin, A.; Wincker, P.; Mauger, S.; Jaillon, O.; Malarme, K.; Weissenbach, J.; Eh			
Genome Res. 11, 731-753, 2001			
A;Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis			
A;Reference number: A86625; MUID:21235186; PMID:11337471			
A;Accession: G86643			
A;Status: preliminary			
A;Molecule type: DNA			
A;Residues: 1-1983 <STO>			
A;Cross-references: GB:AE005176; PID:g12723000; PIDN:AAK04249.1; GSPDB:GN00146			
A;Experimental source: strain IL1403			
C;Genetics:			
A;Gene: ybEF			
Query Match 7.8%; Score 135; DB 2; Length 1983;			
Best Local Similarity 18.7%; Pred. No. 2.7;			
Matches 85; Conservative 50; Mismatches 118; Indels 202; Gaps 17;			

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 23, 2003, 13:59:02 ; Search time 12.3245 Seconds

(without alignments)
1113.936 Million cell updates/sec

Title: US-10-056-052A-4

Perfect score: 1732

Sequence: 1 MVAADAPAAAGTDITNQLTNV.....NGSGSGDGIDKPVVPEQPDE 331

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

1 number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	331	19.1	1018	1	FNBA_STAAU	P14738 staphylococ
2	129	7.4	762	1	SLAP_ACEKI	P22258 acetogenium
3	128	7.4	2021	1	OMPA_RICCN	Q52657 rickettsia
4	127	7.3	642	1	FLID_CAMJE	Q9phw6 campylobact
5	121.5	7.0	2249	1	OMPA_RICRI	P15921 rickettsia
6	119.5	6.9	544	1	GP10_DICDI	Q06885 dictyosteli
7	115.5	6.7	928	1	PM11_CHLPN	O86164 chlamydia p
8	114.5	6.6	1643	1	OMP_B_RICPR	Q53020 r outer mem
9	114	6.6	738	1	YO13_BPL2	P42548 bacterioph
10	113.5	6.6	1228	1	SLAP_BACST	P35825 bacillus st
11	113.5	6.6	1250	1	YFAL_ECOLI	P45508 escherichia
12	113.5	6.6	1902	1	P1P_LACLC	P16271 lactococcus
13	113.5	6.6	1902	1	P2P_LACPA	Q02470 lactobacill
14	113.5	6.6	1902	1	P3P_LACLC	P15292 lactococcus
15	113	6.5	1286	1	AIDA_ECOLI	Q03155 escherichia
16	113	6.5	1394	1	HAP_HAEIN	P45387 haemophilus
17	112	6.5	1902	1	P2P_LACLC	P15293 lactococcus
18	111	6.4	995	1	YIQ9_YEAST	P40442 saccharomyc
19	110.5	6.4	1176	1	SLAP_BACSH	P38537 bacillus sp
20	110.5	6.4	1645	1	OMP_B_RICTY	P96989 r outer mem
21	110	6.4	1183	1	CNA_STAAU	Q53654 staphylococ
22	109.5	6.3	760	1	SIX4_HUMAN	Q9uii6 homo sapien
23	108.5	6.3	2003	1	YDBA_ECOLI	P33666 escherichia
24	108	6.2	827	1	CSG_HALVO	P25062 halobacteri
25	108	6.2	2358	1	YEEJ_ECOLI	P76347 escherichia
26	107.5	6.2	504	1	Y795_METJA	Q58205 methanococc
27	107.5	6.2	1409	1	HAP1_HAEIN	P44596 haemophilus
28	107.5	6.2	1655	1	OMP_B_RICCN	Q9kka3 r outer mem
29	107	6.2	583	1	AP1_KLUJA	P56095 kluyveromyc
30	107	6.2	1698	1	41_DRONE	Q9v8r9 drosophila
31	106.5	6.1	1181	1	ITA2_HUMAN	P17301 homo sapien
32	106	6.1	1561	1	SPAP_STRMU	P23504 streptococc
33	105.5	6.1	402	1	FLGE_SALTY	P16322 salmonella

RESULT 1

ALIGNMENTS

34	105.5	6.1	444	1	SLAP_LACAC	P35829 lactobacill
35	105.5	6.1	544	1	PME3_LYCES	Q96576 lycopersico
36	105.5	6.1	2452	1	RPB1_PLAFD	P14248 plasmodium
37	105	6.1	537	1	TEE6_STRPY	P18481 streptococ
38	105	6.1	655	1	YKDA_MYCCA	P45615 mycoplasma
39	105	6.1	906	1	Y010_CLOAB	Q97n28 clostridium
40	105	6.1	928	1	PMP9_CHLPN	Q92398 chlamydia p
41	105	6.1	1565	1	PAC_STRMU	P11657 streptococ
42	104.5	6.0	448	1	FIBP_ADEP3	Q83457 porcine ade
43	104.5	6.0	870	1	P100_HSV6U	Q00701 human herpe
44	104.5	6.0	936	1	EAE_CITFR	Q07591 citrobacter
45	104	6.0	1656	1	OMP_B_RICJA	Q06653 r outer mem
46	104	6.0	2334	1	WAPA_BACSU	Q07833 bacillus su
47	103.5	6.0	417	1	Y943_METJA	Q58353 methanococc
48	103.5	6.0	643	1	CBAA_BACTI	P21256 bacillus th
49	103.5	6.0	798	1	PBPA_NEICI	O86088 neisseria c
50	103.5	6.0	1654	1	OMP_B_RICRI	Q53047 r outer mem
51	103	5.9	378	1	OMPC_SALTY	O52503 salmonella
52	103	5.9	485	1	TACY_BACCE	Q45105 bacillus ce
53	103	5.9	1723	1	PM20_CHLPN	Q9z812 chlamydia p
54	102.5	5.9	315	1	HEMA_VACCI	P08714 vaccinia vi
55	102.5	5.9	474	1	PEDA_LACHE	Q48558 lactobacill
56	102.5	5.9	1577	1	HLVA_PROMI	P16466 proteus mir
57	102.5	5.9	1848	1	CBPA_CLOCL	P38058 clostridium
58	102	5.9	936	1	PMP7_CHLPN	Q9z898 chlamydia p
59	101.5	5.9	461	1	EPIP_STAEP	P30199 staphylococ
60	101.5	5.9	796	1	CADB_MOUSE	P55288 mus musculu
61	101.5	5.9	799	1	ZFX1_MOUSE	P17011 mus musculu
62	101.5	5.9	839	1	ZFX2_MOUSE	P17012 mus musculu
63	101	5.8	518	1	FLAA_AQUAE	O67803 aquifex aeo
64	101	5.8	553	1	YF61_METJA	Q58956 methanococc
65	101	5.8	983	1	Y644_MYCLE	Q9ccm6 mycobacteri
66	101	5.8	1004	1	SLPO_BACBR	P09333 bacillus br
67	100.5	5.8	315	1	HEMA_VACCC	P20978 vaccinia vi
68	100.5	5.8	373	1	YFJ4_YEAST	P43603 saccharomyc
69	100.5	5.8	466	1	FLID_SALTY	P16328 salmonella
70	100.5	5.8	609	1	LAC1_EMENI	P17489 emericella
71	100.5	5.8	610	1	DNAK_STAAM	Q99tr7 staphylococ
72	100.5	5.8	610	1	DNAK_STAAM	P45554 staphylococ
73	100.5	5.8	634	1	ELM2_ASFPU	P46075 aspergillus
74	100.5	5.8	774	1	LONI_BACSU	P37945 bacillus su
75	100.5	5.8	1341	1	VG37_BPT2	P07067 bacterioph
76	100	5.8	682	1	NISP_LACLA	Q07596 lactococcus
77	100	5.8	3343	1	YOG7_CAELI	P34616 caenorhabdi
78	99.5	5.7	486	1	YAIT_ECOLI	P77199 escherichia
79	99.5	5.7	600	1	DNAK_ERYRH	Q05647 erysipeloth
80	99.5	5.7	917	1	HXA3_HAEIN	P45355 haemophilus
81	99.5	5.7	986	1	GUNZ_CLOSR	P23659 clostridium
82	99.5	5.7	1230	1	ST20_CANAL	Q92212 candida alb
83	99.5	5.7	1279	1	APU_THESA	P36905 t amylopull
84	99.5	5.7	1291	1	VAC4_HELPY	Q48258 helicobacte
85	99.5	5.7	1541	1	IGAI_HAEIN	P42782 haemophilus
86	99.5	5.7	1861	1	APU_THETU	P38536 t amylopull
87	99	5.7	314	1	HEMA_VACCV	Q01218 vaccinia vi
88	99	5.7	928	1	HXA2_HAEIN	P45354 haemophilus
89	99	5.7	1116	1	SLPH_BACBR	P38538 bacillus br
90	99	5.7	1167	1	CLAA_BACTU	P56956 bacillus th
91	99	5.7	1197	1	DPOM_PODAN	Q01529 podospora a
92	99	5.7	2660	1	YEEJ_ECO57	Q8x8v7 escherichia
93	98.5	5.7	484	1	MURE_CLOPE	Q8xj99 clostridium
94	98.5	5.7	493	1	FLIC_SALPA	P06178 salmonella
95	98.5	5.7	548	1	SYFB_METJA	Q58508 methanococc
96	98.5	5.7	634	1	ELM1_ASFPU	P46074 aspergillus
97	98	5.7	387	1	PKNS_MYCGE	P47355 mycoplasma
98	98	5.7	406	1	FLGE_RHIME	Q9x5y0 rhizobium m
99	98	5.7	613	1	CGAA_CLOBI	Q45882 clostridium
100	98	5.7	792	1	CADB_CHICK	O93319 gallus gall

FNBA_STAAU STANDARD; PRT; 1018 AA.
ID FNBA_STAAU STANDARD; PRT; 1018 AA.
AC P14738;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Fibronectin-binding protein precursor (FNBP).
GN FNBA.
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=1280;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NCTC 8325-4; PubMed=2521391;
RX MEDLINE=89098998; PubMed=2521391;
RA Signaes C., Raucchi G., Joensson K., Lindgren P.-E.,
RA Anantharamiah G.M., Hoeoek M., Lindberg M.;
RT "Nucleotide sequence of the gene for a fibronectin-binding protein
from Staphylococcus aureus: use of this peptide sequence in the
synthesis of biologically active peptides";
Proc. Natl. Acad. Sci. U.S.A. 86:699-703(1989).
CC -!- FUNCTION: THE ABILITY OF BACTERIA TO BIND FIBRONECTIN HAS BEEN
PROPOSED AS A VIRULENCE FACTOR ENABLING BACTERIA TO COLONIZE
WOUND TISSUES AND BLOOD CLOTS. BINDING OF PLASMA FIBRONECTIN TO
THE BACTERIAL SURFACE MIGHT BLOCK ADHESION RECEPTORS ON S.AUREUS,
THUS REPRESENTING AN IMPORTANT DEFENSE MECHANISM AGAINST TISSUE
INVASION.
CC -!- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by
an amide bond (Potential).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
CC EMBL; J04151; AAA26632.1; -.
DR InterPro; IPR004237; Fn_bind.
DR InterPro; IPR001899; Gram_pos_anchor.
DR Pfam; PF00746; Gram_pos_anchor; 1.
DR Pfam; PF02986; Fn_bind; 1.
DR TIGRFAMS; TIGR01167; LPXTG_anchor; 1.
DR TIGRFAMS; TIGR01168; YSIRK_signal; 1.
DR PROSITE; PS50847; GRAM_POS_ANCHORING; 1.
KW Cell wall; Peptidoglycan-anchor; Repeat; Signal.
FT SIGNAL 1 36
FT CHAIN 37 985 FIBRONECTIN-BINDING PROTEIN.
FT PROPEP 986 1018 REMOVED BY SORTASE (POTENTIAL).
FT REPEAT 545 574 B-1.
FT REPEAT 575 604 B-2.
FT DOMAIN 745 878 4 X APPROXIMATE TANDEM REPEATS,
FIBRONECTIN-BINDING DOMAIN.
FT REPEAT 745 782 D-1.
FT REPEAT 783 820 D-2.
FT REPEAT 821 859 D-3.
FT REPEAT 860 878 D-4 (INCOMPLETE).
FT DOMAIN 879 948 5 X TANDEM REPEATS, PRO-RICH (WR).
FT REPEAT 879 892 WR1.
FT REPEAT 893 906 WR2.
FT REPEAT 907 920 WR3.
FT REPEAT 921 934 WR4.
FT REPEAT 935 948 WR5.
FT SITE 982 986 LPXTG SORTING SIGNAL (POTENTIAL).
FT MOD_RES 985 985 AMIDE-LINKED TO CELL WALL (POTENTIAL).
SQ SEQUENCE 1018 AA; 111780 MW; 58175E0020E81F1F CRC64;

Query Match 19.1%; Score 331; DB 1; Length 1018;
Best Local Similarity 24.6%; Pred. NO. 3.3e-13;
Matches 82; Conservative 77; Mismatches 146; Indels 28; Gaps 11;

us-10-056-052a-4.rsp

194 GTDVTSKVT-VEIGSIEGHNNTNKVEPHAGQRAVLKYKLKEENGLHQGDYFDTLSNNVN 252
QY 66 LNGVTSTAKVPPIMAGDQVLANGVIDSDGNVIYFTTDYVNTKDDVKATLTMPAYIDPENV 125
Db 253 THGVSTARKVPEIKNGSVVMATGEVLEGGKIRYFTFTNDIEDKVDVTALEINLFIDPKTV 312
QY 126 KKTGNVTLATGIGSTTANKTVLVLDY-EYKGKYNLSIKGTIDQIDKTNNTYRQTIYVNS 184
Db 313 QTNNGQITITSLNEEQTSKELDVYKVDGIGNY-ANLNGSIETEFNKANNRFSHVAFIKPN 371
QY 185 GDNVIAPVLTGNLKPNTDSNALIDQQNTSIKVK-VDNAADLSESYFVN--PENFEDVT 240
Db 372 NGKTSVTVTGTLMKGSNQNG---NQPKVRIFEYLGNNEDIAKSVYANTTDTSKFKEVT 427
QY 241 NSV--NITFPNPNOYKVEFNTDDQITTPYIVVVNGHIDPNKSGDLALRSTLYGYSNII 298
Db 428 SNMSGNLNLQNGSYSLNI-----ENLDKTYVVHYDGEY-LNGTDEVDFRTQMVGHPEQLY 482
QY 299 -!-----WRSMWDNEVAFNNGSGSGDGIDKPVV 325
Db 483 KYVYDRGYTLTWDNGLVLYSNKANGNEKNGPII 515

RESULT 2
SLAP_ACEKI STANDARD; PRT; 762 AA.
AC P22258;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Cell surface protein precursor (S-layer protein).
OS Acetogenium kivui.
OC Bacteria; Firmicutes; Clostridia; Thermoanaerobacteriales;
OC Thermoanaerobacteriaceae; Thermoanaerobacter.
OX NCBI_TaxID=2325;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC STRAIN=DSM 2030;
RX MEDLINE=90036724; Pubmed=2681162;
RA Peters J., Peters M., Lottspeich F., Baumeister W.;
RT "S-layer protein gene of Acetogenium kivui: cloning and expression in
Escherichia coli and determination of the nucleotide sequence.";
RL J. Bacteriol. 171:6307-6315(1989).
RN [2]
RP PARTIAL SEQUENCE, AND CARBOHYDRATE-LINKAGE SITES.
RC STRAIN=DSM 2030;
RX MEDLINE=92281680; Pubmed=1596358;
RA Peters J., Rudolf S., Oschkinat H., Mengele R., Sumper M.,
RA Kellermann J., Lottspeich F., Baumeister W.;
RT "Evidence for tyrosine-linked glycosaminoglycan in a bacterial
surface protein.";
RL Biol. Chem. Hoppe-Seyler 373:171-176(1992).
RN [3]
RP DOMAINS.
RX MEDLINE=94156823; Pubmed=8113161;
RA Lupas A., Engelhardt H., Peters J., Santarius U., Volker S.,
RA Baumeister W.;
RT "Domain structure of the Acetogenium kivui surface layer revealed by
electron crystallography and sequence analysis.";
RL J. Bacteriol. 176:1224-1233(1994).
CC -!- FUNCTION: THE S-LAYER IS A PARACRYSTALLINE MONO-LAYERED ASSEMBLY
OF PROTEINS WHICH COAT THE SURFACE OF BACTERIA.
CC -!- SUBCELLULAR LOCATION: CELL WALL. THIS BACTERIUM IS COVERED BY A
S-LAYER WITH HEXAGONAL SYMMETRY.
CC -!- PTM: THE CARBOHYDRATE CONTENT OF THIS PROTEIN IS ABOUT 8% WHICH
CORRESPOND TO ABOUT 40 TO 50 SUGAR MOLECULES PER MONOMER. O-LINKED
GLYCANS CONSIST OF GLC, GALNAC AND GLCNAC.
CC -!- SIMILARITY: CONTAINS 3 S-LAYER HOMOLGY (SLH) DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entitles requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL; M31069; AAA21930.1; -.
DR PIR; A34355; A34355.
DR InterPro; IPR001119; SLH.
DR Pfam; PF00395; SLH; 2.
DR PROSITE; PS01072; SLH_DOMAIN; 2.
KW Signal; Glycoprotein; Repeat; Cell wall; S-layer.
FT SIGNAL 1 26
FT CHAIN 27 762 CELL SURFACE PROTEIN.
FT DOMAIN 30 94 SLH 1.
FT DOMAIN 95 145 SLH 2.
FT DOMAIN 146 204 SLH 3.
FT DOMAIN 473 479 SER/THR-RICH.
FT DOMAIN 625 630 SER/THR-RICH.
FT CARBOHYD 297 297 O-LINKED (GLC. . .).
FT CARBOHYD 516 516 O-LINKED (GLC. . .).
FT CARBOHYD 520 520 O-LINKED (GLC. . .).
FT CARBOHYD 632 632 O-LINKED (GLC. . .).
SQ SEQUENCE 762 AA; 82785 MW; 34EC9C784DECA67E CRC64;

Query Match 7.4%; Score 129; DB 1; Length 762;
Best Local Similarity 21.6%; Pred. No. 0.72;
Matches 79; Conservative 54; Mismatches 113; Indels 120; Gaps 19;
Qy 15 NQLTNVTVGIDS-GTTVYPHQAGYVKLNFGFSVPNSAVKGDTEFKITVPKELN----LNG- 68
Db 280 NDVVSFTEGQDSVGTIVY-----KNDNKNKTAIKVDDNAYVLYNGY 319
Qy 69 VTSTAKVP-----PIMAGDOVLANGVIDSDG---NVIYTFDYVN--TKDDVVKATLTM 116
Db 320 LTKVSKVTVKEGAETVIINNLYLVNGSYDNSTIVYNDVQSGDKYLNKRDSDNYELKGTIVTV 379
Qy 117 PAVIDPENVKKTGNVTLATGIGSTTANKTVLVVDYEKYPYNLS-----IK 162
Db 380 -----TGAVSKVTDI---KAN---DYIYKQYDVNGVNGVTVIYVVRNQVT 420
Qy 163 GTIDQIDKTNNTYRQTIYVNPNGDNVAPVLGTN-----LKPNTDSNALIDQONTSTIKVYK 218
Db 421 GTVTEKSVSGSTYKASI-----DNVSYTVADNNVWVNOLEPGKKVTVLNKNQNVIGI-- 472
Qy 219 VDNAADLSESYFVNPENFEDVTVNSVNITFPNPNQYKVFNTPD-----DOI 264
Db 473 ----SSTTTTAVNYAIFKEKSDPFTAWFA----KVKLLLPDAAEKVFDVAVSDYVDKV 523
265 TTPYIVVWNGHIDPNK-----GDALRSTLYGYNLSIWRSMS----WDNEVAFNNG 313
Db 524 NLAEGTIVTYVDANGKLNLDIQRANDQPFSSASYKADAKVLTGSTTYIITDNTVLLNNT 583
Qy 314 SSGSDG 319
Db 584 S---DG 586

RESULT 3
OMPA_RICCN STANDARD; PRT: 2021 AA.
AC Q52657; P95591; P95592; P95593; P95594; Q52667; Q52668; Q52669;
AC Q52670; Q52674;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Outer membrane protein A precursor (190 kDa antigen) (Cell surface
DE antigen) (rOmpA) (rOmp A).
GN OMPA OR RC1273.
OS Rickettsia conorii.
OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
OC Rickettsiaceae; Rickettsieae; Rickettsia.
OX NCBI_TaxID=781;
RN [1]

RP SEQUENCE FROM N.A.
RC STRAIN=Malish 7;
RX MEDLINE=94171067; PubMed=8125327;
RA Crocquet-Valdes P.A., Weiss K., Walker D.H.;
RT "Sequence analysis of the 190-kDa antigen-encoding gene of Rickettsia
RL conorii (Malish 7 strain).";
RN Gene 140:115-119(1994).
RP SEQUENCE FROM N.A.
RC STRAIN=Malish 7;
RX MEDLINE=21442074; PubMed=11557893;
RA Ogata H., Audic S., Renesto-Audiffren P., Fournier P.-E., Barbe V.,
RA Samson D., Roux V., Cossart P., Weissenbach J., Claverie J.-M.,
RA Raoult D.;
RT "Mechanisms of evolution in Rickettsia conorii and R. prowazekii.";
RL Science 293:2093-2098(2001).
RN [3]
RP SEQUENCE OF 8-204 FROM N.A.
RC STRAIN=Indian tick typhus, M1, Malish 7, and Moroccan;
RX MEDLINE=97015921; PubMed=8862558;
RA Roux V., Fournier P.E., Raoult D.;
RT "Differentiation of spotted fever group rickettsiae by sequencing and
analysis of restriction fragment length polymorphism of PCR-amplified
RT DNA of the gene encoding the protein rOmpA.";
RL J. Clin. Microbiol. 34:2058-2065(1996).
RN [4]
RP SEQUENCE OF 953-2012 FROM N.A.
RC STRAIN=Indian tick typhus, M1, Malish 7, and Moroccan;
RA Raoult D., Fournier P.E., Roux V.;
RT "Phylogenetic analysis of spotted fever group rickettsiae by study
of the outer surface protein rOmpA.";
RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: ELICITS PROTECTIVE IMMUNITY (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: CELL WALL. THIS BACTERIUM IS COVERED BY A
CC S-LAYER WITH HEXAGONAL SYMMETRY.
CC -!- PTM: GLYCOSYLATED (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE RICKETTSIAE OMPA/OMPB FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U01028; AAA17405.1; -.
DR EMBL; AE008674; AAL03811.1; -.
DR EMBL; U43794; AAB49549.1; -.
DR EMBL; U43798; AAB49550.1; -.
DR EMBL; U43806; AAB49551.1; -.
DR EMBL; U45244; AAB49566.1; -.
DR EMBL; U46918; AAB49566.1; -.
DR EMBL; U83440; AAC35176.1; -.
DR EMBL; U83443; AAC35179.1; -.
DR EMBL; U83448; AAC35184.1; -.
DR EMBL; U83453; AAC35189.1; -.
DR InterPro; IPR003858; rOmpA_rOmpB.
DR Pfam; PF02708; rOmpA_rOmpB; 1.
KW Antigen; Repeat; Signal; Cell wall; S-layer; Glycoprotein;
KW Complete proteome.
FT SIGNAL 1 38
FT CHAIN 39 2021 POTENTIAL.
FT DOMAIN 238 946 OUTER MEMBRANE PROTEIN A.
FT DOMAIN 1424 1528 THR-RICH.
FT VARIANT 60 60 THR-RICH.
FT VARIANT 76 76 R -> NN (IN STRAIN INDIAN TICK TYPHUS).
FT VARIANT 86 137 R -> H (IN STRAIN INDIAN TICK TYPHUS).
FT VARIANT 126 133 MISSING (IN STRAIN MOROCCAN).
FT VARIANT 953 954 VT -> II (IN STRAIN INDIAN TICK TYPHUS).
FT VARIANT 1245 1245 D -> A (IN STRAINS INDIAN TICK TYPHUS, M1
AND MOROCCAN).
FT VARIANT 1308 1308 N -> H (IN STRAIN MOROCCAN).

FT VARIANT 1877 1877 M -> I (IN STRAIN INDIAN TICK TYPHUS).
FT CONFLICT 10 10 Q -> K (IN REF. 1).
FT CONFLICT 92 92 I -> V (IN REF. 1).
FT CONFLICT 126 126 V -> I (IN REF. 1).
FT CONFLICT 137 137 T -> N (IN REF. 1).
FT CONFLICT 157 157 G -> D (IN REF. 1).
FT CONFLICT 368 369 IS -> VN (IN REF. 1).
FT CONFLICT 374 388 KATLGGAIIKATTTK -> LLQVGGVVKANTIN (IN REF. 1).
FT CONFLICT 640 640 N -> D (IN REF. 1).
FT CONFLICT 669 669 V -> I (IN REF. 1).
FT CONFLICT 793 793 N -> D (IN REF. 1).
FT CONFLICT 803 804 VN -> IS (IN REF. 1).
FT CONFLICT 809 823 LLRVQGGVVKSNIN -> KATLGGAIIKATTTK (IN REF. 1).
FT CONFLICT 898 898 D -> Y (IN REF. 1).
FT CONFLICT 908 908 P -> N (IN REF. 1).
FT CONFLICT 985 985 N -> K (IN REF. 1).
FT CONFLICT 1009 1009 L -> S (IN REF. 1).
FT CONFLICT 1013 1013 Y -> S (IN REF. 1).
FT CONFLICT 1182 1182 K -> Q (IN REF. 1).
FT CONFLICT 1314 1314 N -> Y (IN REF. 4).
FT CONFLICT 1451 1451 H -> N (IN REF. 1).
FT CONFLICT 1624 1624 G -> D (IN REF. 1).
FT CONFLICT 1628 1628 E -> G (IN REF. 1).
FT CONFLICT 1872 1872 A -> V (IN REF. 1).
FT CONFLICT 1875 1875 T -> P (IN REF. 1).
FT CONFLICT 1878 1879 MS -> LP (IN REF. 1).
FT CONFLICT 1936 1936 E -> A (IN REF. 1).
FT CONFLICT 1965 1970 MTAPLP -> ITPPLS (IN REF. 1).
FT CONFLICT 1997 1997 G -> R (IN REF. 1).
SQ SEQUENCE 2021 AA; 203328 MW; 327FC42D7CB24668 CRC64;

Query Match 7.4%; Score 128; DB 1; Length 2021;
Best Local Similarity 25.1%; Pred. No. 2.9;
Matches 90; Conservative 31; Mismatches 106; Indels 132; Gaps 20;
QY 8 AAGTDITN----QLTNV---TVGIDSGTTVYPHQAGYVKLNYGFS----- 45
Db 672 ATTKLTNAASVLTLTNVNAVLTGAIDNTTGV--DNVGVNLNGALSQVTGNIGNTNALA 729
QY 46 -----VPNSAVKGDTEFKITVPKELNNGVTSFAKVPPIPMAGDQVLANGVIDSDGN 95
Db 730 TISVGAGKATLGGAVIKATTKLTD---NASAVTFT-----NPVVVTGAIDNTGN 776
QY 96 V---IYTFDYVNTKDDVKATLTMPAYIDPENVKKTGNVTLATGIGSTTANKTV----- 146
Db 777 ANNGIATF-----TGDSVTVTGNIGNTNALATVNVGAGL 809
QY 147 -----LVDYKEYGKFYN-LSIKGTIDQIDKTNN-----TYRQTIYVNPSPGD 186
Db 810 LRVOGGVVKSNITINLTDNASAVTFTNPVVVTGAIDNTGNANNGIVTFTGDSVTGNIGNT 869
QY 187 NVIAPVLTGNLKPNTDSNALIDQQNTSIKVKVDNAADLSESYFVNPFENFEDVTNSVNIT 246
Db 870 NALATISVGAGKA-TLGGAIKATTTKL---TDNASAVT---FTNP---VVVTGAIDNT 918
QY 247 FPNPNQYKVEFNTPDQDIT-----TPYIVVNVN-----GHIDPNSKGDALRSTL 290
Db 919 -GNANNGIVTF-TGDSVTVTGNIGNTNALATVNVGAGVTLOAGGSLDANNI-DEGARSTL 974

RESULT 4
FLID_CAMJE STANDARD; PRT; 642 AA.
AC Q9PHW6;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Flagellar hook-associated protein 2 (HAP2) (Filament cap protein)
DE (Flagellar cap protein).
GN FLID OR CJ0548.
OS Campylobacter jejuni.

OC Bacteria; Proteobacteria; epsilon subdivision; Campylobacter group.
OC Campylobacter.
OX NCBI_TaxID=197;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NCTC 11168;
RX MEDLINE=20150912; PubMed=10688204;
RA Parkhill J., Wren B.W., Mungall K., Ketley J.M., Churcher C.,
RA Basham D., Chillingworth T., Davies R.M., Feltwell T., Holtroyd S.,
RA Jagels K., Karlyshev A.V., Moule S., Pallen M.J., Penn C.W.,
RA Quail M.A., Rajandream M.A., Rutherford K.M., van Vliet A.H.M.,
RA Whitehead S., Barrell B.G.;
RT "The genome sequence of the food-borne pathogen Campylobacter jejuni
RT reveals hypervariable sequences.";
RL Nature 403:665-668(2000).
CC -!- FUNCTION: REQUIRED FOR THE MORPHOGENESIS AND FOR THE ELONGATION OF
CC THE FLAGELLAR FILAMENT BY FACILITATING POLYMERIZATION OF THE
CC FLAGELLIN MONOMERS AT THE TIP OF GROWING FILAMENT. FORMS A CAPPING
CC STRUCTURE, WHICH PREVENTS FLAGELLIN SUBUNITS (TRANSPORTED THROUGH
CC THE CENTRAL CHANNEL OF THE FLAGELLUM) FROM LEAKING OUT WITHOUT
CC POLYMERIZATION AT THE DISTAL END (BY SIMILARITY).
CC -!- SUBUNIT: HOMOPENTAMER (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Flagellar.
CC -!- SIMILARITY: BELONGS TO THE FLID FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AL139075; CAB75184.1; -
DR InterPro; IPR003481; Flid.
DR Pfam; PF02465; Flid; 1.
KW Flagella; Coiled coil; Complete proteome.
FT DOMAIN 614 641 COILED COIL (POTENTIAL).
SQ SEQUENCE 642 AA; 69775 MW; 40FF20D41668EFA3 CRC64;

Query Match 7.3%; Score 127; DB 1; Length 642;
Best Local Similarity 20.9%; Pred. No. 0.76;
Matches 86; Conservative 48; Mismatches 136; Indels 142; Gaps 19;
QY 21 TVGIDSGTTVYPHQAGYVKLNYG---FVSPNSAVKGDTEFKITV-PKELNNGVTSFAKV 75
Db 146 TVTVDKNTT-YRLADKINEASGEIVAKIVNTGKGPYRLTLTSKETGEDSAIS---- 200
QY 76 PPIMAGDQVLANGVIDSDGNVIYTFDYVNTKDDVKATLTMPAYIDPENVKK----- 127
Db 201 --FYAGKKD-SNGKYQKDINAEIF-----DDLGLDVSASIDPDKKKGYGKIDAS 250
QY 128 -----TGNVTLATGIGST-TANKTVLVDYKEYGKFYNLSIKGTID 166
Db 251 LHIQTAQNAEFTLDGIKMFRSSNTVTDLGVGMTLTLNKTGEINFVQDQDFEGVT-KAMQD 309
QY 167 QIDKTN----NTYRQTIY-----VNPSGDNVIAPVLTGNLKPNTDSNALI 207
Db 310 LVDAYNDLVTLNLAATDYNSETGKTGTLQGISEVNSIRSSILADLFDSDVVDGTTEDANG 369
QY 208 DQNTSIKV-----YKVDNAADLSESYFVNPFENFEDVT---- 240
Db 370 NKVNTKVMLSMQDFGLSLNDAGTSLFSDSSKFEQVKVEDPDSTESFFSNITKYEDINHTE 429
QY 241 -----NSVNITFPNPQYKVEFNTPDQDIT-----TPYIVV 271
Db 430 VIKTGSLSKYLNSNGGNTNGLEF-KPGDFTIVFNNTQYDLSKNSDGTNFKLTGKTEELL 488
QY 272 VN--GHIDPNSKGDALRSTLYGYNLSIIWRSMSWDNEVAFN---NGSGSGD 318
Db 489 QNLANHI--NSKGIEGLKVKVESYNQN-----NVTGRLNFSGSGSSD 529


```
RESULT 8
OMP_RICPR
ID OMPB_RICPR STANDARD; PRT; 1643 AA.
AC Q53020; Q92CM0;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Outer membrane protein B precursor (168 kDa surface-layer protein)
DE (Surface protein antigen) (Cell surface antigen 5) (Sca5) (rOmpB)
DE (rOmp B) [Contains: 120 kDa surface-exposed protein (Surface protein
DE antigen) (120 kDa outer membrane protein ompB); 32 kDa beta peptide].
GN OMPB OR SPAP OR SPA OR RP704.
OS Rickettsia prowazekii.
OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
OC Rickettsiaceae; Rickettsiae; Rickettsia.
OX NCBI_TaxID=782;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
MEDLINE=91045972; PubMed=2122457;
MEDLINE=91045972; PubMed=2122457;
Carl M., Dobson M.E., Ching W.M., Dasch G.A.;
"Characterization of the gene encoding the protective paracrystalline-
surface-layer protein of Rickettsia prowazekii: presence of a
truncated identical homolog in Rickettsia typhi.";
Proc. Natl. Acad. Sci. U.S.A. 87:8237-8241(1990).
[2]
RP SEQUENCE FROM N.A.
STRAIN=Breinl;
RA Moron C.G., Yu X.J., Walker D.H.;
"Sequence analysis of ompB of Rickettsia prowazekii.";
Submited (JUN-1999) to the EMBL/GenBank/DBJ databases.
[3]
RP SEQUENCE FROM N.A.
STRAIN=Breinl;
RA Andersson S.G.E., Zomorodipour A., Andersson J.O.,
Sicheritz-Ponten T., Alsmark U.C.M., Podowski R.M., Naeslund A.K.,
Eriksson A.-S., Winkler H.H., Kurland C.G.;
"The genome sequence of Rickettsia prowazekii and the origin of
mitochondria.";
Nature 396:133-140(1998).
[4]
RP PARTIAL SEQUENCE.
STRAIN=Breinl;
RA Ching W.M., Carl M., Dasch G.A.;
"Mapping of monoclonal antibody binding sites on CNBr fragments of
the S-layer protein antigens of Rickettsia typhi and Rickettsia
prowazekii.";
Mol. Immunol. 29:95-105(1992).
[5]
RP CLEAVAGE SITE.
MEDLINE=92104668; PubMed=1729180;
Hackstadt T., Messer R., Cieplak W. Jr., Peacock M.G.;
"Evidence for proteolytic cleavage of the 120-kilodalton outer
membrane protein of rickettsiae: identification of an avirulent mutant
deficient in processing.";
Infect. Immun. 60:159-165(1992).
-1- FUNCTION: THE 120 kDa SURFACE-EXPOSED PROTEIN IS A MAJOR
STRUCTURAL PROTEIN WHICH MAY PLAY A ROLE AS A RICKETTSIAL
VIRULENCE FACTOR AND/OR IMMUNOGEN DURING INFECTION.
-1- FUNCTION: THE 32 kDa BETA PEPTIDE MAY SERVE AS A MEMBRANE ANCHOR.
-1- SUBCELLULAR LOCATION: CELL WALL. THIS BACTERIUM IS COVERED BY A S-
LAYER WITH HEXAGONAL SYMMETRY.
-1- SIMILARITY: BELONGS TO THE RICKETTSIAE OMPA/OMPB FAMILY.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/)
```

```
or send an email to license@isb-sib.ch.
-----
CC EMBL; M37647; AAA26390.1; ALT_INIT.
CC EMBL; AF161079; AAD42234.1; -.
DR EMBL; AJ235273; CAA15140.1; -.
DR InterPro; IPR003858; rOmpA_rOmpB.
DR Pfam; PF02708; rOmpA_rOmpB; 1.
KW Antigen; S-layer; Cell wall; Complete proteome.
FT CHAIN 1 1328 120 KDA SURFACE-EXPOSED PROTEIN.
FT CHAIN 1329 1643 32 KDA BETA PEPTIDE.
FT VARIANT 257 257 V -> A (IN STRAIN BREINL).
FT VARIANT 1010 1010 Y -> D (IN STRAIN BREINL).
FT VARIANT 1450 1450 A -> S (IN STRAIN BREINL).
FT CONFLICT 178 179 AA -> VC (IN REF. 1).
FT CONFLICT 191 201 TTQEAPLTLGA -> INSRSSSYHLVS (IN REF. 1).
FT CONFLICT 212 212 T -> I (IN REF. 1).
FT CONFLICT 313 313 Q -> L (IN REF. 1).
FT CONFLICT 1104 1104 D -> G (IN REF. 2).
FT CONFLICT 1123 1123 T -> S (IN REF. 2).
SQ SEQUENCE 1643 AA; 169854 MW; 735FDF392E6346CC CRC64;

Query Match 6.6%; Score 114.5; DB 1; Length 1643;
Best Local Similarity 21.0%; Pred. No. 15;
Matches 80; Conservative 48; Mismatches 140; Indels 113; Gaps 17;

QY 8 AAGTDITNQLTNVT---VGIDSGTTVYPHQAGYVVKLVNGFESVPNSAVKG---DTFKITVP 61
Db 127 AAG-----KILNITQGQITVQVEASNTINAQNALTKVHGGAANANDLSGLGSITF-AAAP 180

QY 62 KELNLNGVTSTAKVPPIMAGDQVLANGVIDSGNVITYFTD-YVNTKDDVKATL----- 114
Db 181 SVLEFNLIPTTQEAPLTLG---ANSKIVNGGNGTLNTNGFIQVSDNTFAGIKTINID 236

QY 115 -----TMPAYIDPENVKKTGNVTLATGIGSTTANKTVLVLDYKYGKFNLSIKGTID 166
Db 237 DCQGLMFNSTPDAANTLNQVGGNTINFNGIDGT--GKLVLV--SKNGAATEFNVGTGL- 291

QY 167 QIDKTNNTYRQTIYVNPSPGDNVIAPVLTGNLKPNTDSNA--LIDQQNTSIKVKYKVDNAAD 224
Db 292 -----CGN-----LKGIIELNTAAVAGKLLISQGGAAANAVIGTDNGAG 328

QY 225 LSESYFVNPE-----NFEDVTNSV 243
Db 329 RAAGFIVSDNGNAATISQVYAKNMVIOQSANAGGVTFEHIQVGLGGTTNFKTADSKV 388

QY 244 NITFPNPQYKVEFTPDQITTPYIVVVNGHI--DPNSKGDALRSTLYGYNLSIIWRS 301
Db 389 IIT-ENSFNGSTNFGNLDITQIVVPDTKILKGNFIDGVKNNGNTA---GVITFNANGALVS 444

QY 302 MSWDNEVAFNN-----GSGSG 317
Db 445 ASTDPNIAVTNINAIEAEGAG 465

RESULT 9
YO13_BPL2
ID YO13_BPL2 STANDARD; PRT; 738 AA.
AC P42548;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE Hypothetical 81.3 kDa protein precursor (ORF13).
OS Bacteriophage L2.
OC Viruses; dsDNA viruses, no RNA stage; Plasmaviridae; Plasmavirus.
OX NCBI_TaxID=46014;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94215869; Pubmed=8163159;
RA Maniloff J., Kampo G.J., Dascher C.C.;
RT "Sequence analysis of a unique temperature phage: mycoplasma virus
RL L2.";
RL Gene 141:1-8(1994).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
```


Best Local Similarity 23.2%; Pred. No. 21;
Matches 76; Conservative 35; Mismatches 104; Indels 113; Gaps 17;

QY 3 AADAPAAAGTDITNQ-----LTNV-----TVGIDSGT-----TVYPHQAGYVKL- 40
Db 506 AQAAGAAGLIIVNDGTATPLTSIRLTTFPTFGLSSKTGQKLVWDVTAHPDDSLGVKIA 565
QY 41 -----NYGFSVPNSAVKGDTEKITVP-----KELNLNGVTS---TAK 74
Db 566 LTLPLNQKYTEDKMSDFTSYG-PVSNLSFKPD---ITAPGGNIWSTQNNNGYTNMSGTSM 621
QY 75 VPPIMAGDQVLANGVIDSDGNVIYTFDYVNTKDDVKATLTMPAYIDPENVKKTGNVTLA 134
Db 622 ASPEIAGSQALLKQALNNKNPFY--ADYKQLKG---TALT-----DFLKTVMEMNTA 668
QY 135 TGIGSTTANKTVLVDYEKYGKFNLSIKGFIDQIDK-----TNN 173
Db 669 QPINDINYN-NVIVSPRROGAGL-VDVKAAIDALEKNPSTVVAENGYPAVELKDFSTDK 726
QY 174 TYRQTIYVNPNSGDNIAPVLTGNLKPNTDSNAL----IDQNTSIKVKYKVDNAADLSESY 229
Db 727 TEKLTTF-----TNRTTHELTQYQMSNTDTNAVYTSATDPNSGVLYDKKIDGAA----- 774
QY 230 FVNPENFEDVTSVNIITFPNPQYKVEF 257
Db 775 -----IKAGSDITVPAGKTAQIEF 793

RESULT 14
P3P_LACLC STANDARD; PRT; 1902 AA.
AC P15292;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE PIII-type proteinase precursor (EC 3.4.21.96) (Lactocepin) (Cell wall-associated serine proteinase).
GN PRTP.
OS Lactococcus lactis (subsp. cremoris) (Streptococcus cremoris).
OG Plasmid.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.
OX NCBI_TaxID=1359;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 188-197.
RC STRAIN=SK11;
RX MEDLINE=89340435; PubMed=2760036;
RA Vos P., Simons G., Slezten R.J., de Vos W.M.;
"Primary structure and organization of the gene for a procaryotic, cell envelope-located serine proteinase."
J. Biol. Chem. 264:13579-13585(1989).
CC -!- FUNCTION: PROTEASE WHICH BREAKS DOWN MILK PROTEINS DURING THE GROWTH OF THE BACTERIA ON MILK.
CC -!- CATALYTIC ACTIVITY: Endopeptidase activity with very broad specificity, although some subsite preference have been noted, e.g. large hydrophobic residues in the P1 and P4 positions, and Pro in the P2 position. Best known for its action on caseins, although it has been shown to hydrolyse hemoglobin and oxidized insulin B-chain.
CC -!- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by an amide bond (Potential).
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC -----
DR EMBL: J04962; AAA03533.1; ALT_SEQ.
DR PIR: A32634; A32634.
DR HSSP: P00782; 2SBT.

MEROPS; S08.019; .
DR InterPro; IPR001899; Gram_pos_anchor.
DR InterPro; IPR003137; PA.
DR InterPro; IPR000209; Peptidase_S8.
DR Pfam; PF00082; Peptidase_S8; 3.
DR Pfam; PF00746; Gram_pos_anchor; 1.
DR Pfam; PF02225; PA; 1.
DR PRINTS; PR00723; SUBTILISIN.
DR TIGRFAMS; TIGR01167; LPXTG_anchor; 1.
DR PROSITE; PS50847; GRAM_POS_ANCHORING; 1.
DR PROSITE; PS00136; SUBTILASE_ASP; 1.
DR PROSITE; PS00137; SUBTILASE_HIS; 1.
DR PROSITE; PS00138; SUBTILASE_SER; 1.
KW Hydrolase; Serine protease; Cell wall; peptidoglycan-anchor; zymogen;
KW Signal; Plasmid.
FT SIGNAL 1 33
FT PROPEP 34 187
FT CHAIN 188 1870
FT PROPEP 1871 1902 PIII-TYPE PROTEINASE.
FT ACT_SITE 217 217 REMOVED BY SORTASE (POTENTIAL).
FT ACT_SITE 281 281 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 620 620 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT SITE 1867 1871 LPXTG SORTING SIGNAL (POTENTIAL).
FT MOD_RES 1870 1870 AMIDE-LINKED TO CELL WALL (POTENTIAL).
SQ SEQUENCE 1902 AA; 200550 MW; 87CECBAA9345F9D3 CRC64;

Query Match 6.6%; Score 113.5; DB 1; Length 1902;
Best Local Similarity 21.0%; Pred. No. 21;
Matches 70; Conservative 38; Mismatches 101; Indels 125; Gaps 15;

QY 3 AADAPAAAGTDITN-----QLTNV-----TVGIDSGT-----TVYPHQAGYVKLN 41
Db 506 AQAAGAAGLIIVNTDGTATPMTSIALTTTFPTFGLSSVTGQKLVWDVTAHPDDSLGVKIT 565
QY 42 YGFSVPNSAVKGD-----TEK---ITVP-----KELNLNGVTS---TAKVPP 77
Db 566 LAM-LPNQKYTEDKMSDFTSYGVPVSNLSFKPDITAPGGNIWSTQNNNGYTNMSGTSMASP 624
QY 78 IMAGDQVLANGVIDSDGNVIYTF-----TDYVNTKDDVKATLTMPAYIDPENVKKT 128
Db 625 FIAGSQALLKQALNNKNPNPFYAYYKQLKGALTDFLKT-----VEMNTAQPI 671
QY 129 GNVTLATGIGSTANKTVLVDYEKYGKFNLSIKGTIDQIDK----- 170
Db 672 NDINYNVIVSPRROGAGLVD-----VKAIDALEKNPSTVVAENGYPAVELKD 720
QY 171 ---TNNTYRQTIYVNPNSGDNIAPVLTGNLKPNTDSNAL----IDQNTSIKVKYKVDNAA 223
Db 721 FTSTDKTEKLTTF-----TNRTTHELTQYQMSNTDTNAVYTSATDPNSGVLYDKKIDGAA 774
QY 224 DLSESYFVNPENFEDVTSVNIITFPNPQYKVEF 257
Db 775 -----IKAGSNITVPAGKTAQIEF 793

RESULT 15
AIDA_ECOLI STANDARD; PRT; 1286 AA.
AC Q03155;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Adhesin aida-I precursor.
GN AIDA-I.
OS Escherichia coli.
OG Plasmid pIB6.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 50-56.
RC STRAIN=O126:H27 / 2787;
RX MEDLINE=92326638; PubMed=1625582;


```
RESULT 17
P2P_LACLC
ID P2P_LACLC STANDARD; PRT; 1902 AA.
.AC P15293;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
.DE PII-type proteinase precursor (EC 3.4.21.96) (Lactocepina) (Cell wall-
.DE associated serine proteinase) (LP151).
GN PRT.
OS Lactococcus lactis (subsp. cremoris) (Streptococcus cremoris).
OG Plasmid pLP763.
.OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.
.OX NCBI_TaxID=1359;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NCDO 763;
MEDLINE=89313288; PubMed=2501630;
KI Kiwaki M., Ikemura H., Shimizu-Kadota M., Hirashima A.;
RT "Molecular characterization of a cell wall-associated proteinase gene
RL from Streptococcus lactis NCDO763.";
RL Mol. Microbiol. 3:359-369(1989).
CC -!- FUNCTION: PROTEASE WHICH BREAKS DOWN MILK PROTEINS DURING THE
CC GROWTH OF THE BACTERIA ON MILK.
CC -!- CATALYTIC ACTIVITY: Endopeptidase activity with very broad
CC specificity, although some subsite preference have been noted,
CC e.g. large hydrophobic residues in the P1 and P4 positions, and
CC Pro in the P2 position. Best known for its action on caseins,
CC although it has been shown to hydrolyse hemoglobin and oxidized
CC insulin B-chain.
CC -!- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by
CC an amide bond (Potential).
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X14130; CAA32350.1; -.
CC PIR; S06997; S06997.
CC HSSP; P00782; 2SBT.
CC -----
CC MEROPS; S08.019; -.
CC InterPro; IPR001899; Gram_pos_anchor.
CC InterPro; IPR00137; PA.
CC InterPro; IPR00209; Peptidase_S8.
CC Pfam; PF00082; Peptidase_S8; 3.
CC Pfam; PF00746; Gram_pos_anchor; 1.
CC Pfam; PF02225; PA; 1.
CC PRINTS; PR00723; SUBTILISIN.
CC TIGRFAMs; TIGR01167; LPXTG_anchor; 1.
CC PROSITE; PS50847; GRAM_POS_ANCHORING; 1.
CC PROSITE; PS00136; SUBTILASE_ASP; 1.
CC PROSITE; PS00137; SUBTILASE_HIS; 1.
CC PROSITE; PS00138; SUBTILASE_SER; 1.
CC Hydrolase; Serine protease; Cell wall; Peptidoglycan-anchor; Zymogen;
CC Signal; Plasmid.
FT SIGNAL 1 33 POTENTIAL.
FT PROPEP 34 187 POTENTIAL.
FT CHAIN 188 1870 PII-TYPE PROTEINASE.
FT PROPEP 1871 1902 REMOVED BY SORTASE (POTENTIAL).
FT ACT_SITE 217 217 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 281 281 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 620 620 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT SITE 1867 1871 LPXTG SORTING SIGNAL (POTENTIAL).
FT MOD_RES 1870 1870 AMIDE-LINKED TO CELL WALL (POTENTIAL).
SQ SEQUENCE 1902 AA; 200139 MW; 4B8D8B844D88CDF7 CRC64;
```

```
Query Match 6.5%; Score 112; DB 1; Length 1902;
Best Local Similarity 21.1%; Pred. No. 26;
Matches 60; Conservative 37; Mismatches 90; Indels 98; Gaps 13;

QY 21 TVGIDSGTTVYPHQAGYVK-----LNYGFSVPNSAVKGDFTFKITVP-----KELNLNGV 69
   ::::: |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::|
Db 559 SLGVKIALTLVPNQK-YTEDKMSDFTSYG-PVSNLSFKPD---ITAPGGNIWSTQNNNGY 613
   ::::: |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::|
QY 70 TS---TAKVPPIMAGDQVLANGVIDSDGNVIYTF-----TDYVNTKDDVKATLTMP 117
   ::::: |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::|
Db 614 TNMSGTSMASPFIAQSQALLKQALNNKNNPFYAYYKQLKGTALTDLFKT----- 662
   ::::: |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::|
QY 118 AYIDPENVKKTGNVTLATGIGSTTANKTVLVDYEKYGKFYNLSIKGTIDQIDK----- 170
   ::::: |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::|
Db 663 --VEMNTAQPININYNVIVSPRQAGLVD-----VKAIDALEKNPSTVVA 709
   ::::: |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::|
QY 171 -----TNNTYRQTIYVNPSTGDNVIAPVLTGNLKPNTDSNAL-----IDQNT 212
   ::::: |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::|
Db 710 ENGYPAVELKDETSTDKTKLTF-----TNRTTHTLYQMDSNSTDTNAVYTSATDPNSG 763
   ::::: |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::|
QY 213 SIKVYKVDNAADLSESYFVNPFENFEDVTNSVNITFPNPNOYKVEF 257
   ::::: |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::|
Db 764 VLYDKKIDGAA-----IKAGSNITVPAGKTAQIEF 793
   ::::: |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::|

RESULT 18
YIQ9_YEAST
ID YIQ9_YEAST STANDARD; PRT; 995 AA.
AC P40442;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical 99.7 kDa protein in SDL1 5' region precursor.
GN YIL169C OR YI9402.07C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288c / AB972;
RA Barrell B.G., Badcock K., Bankier A.T., Bowman S., Brown D.,
RA Churcher C.M., Connor R., Copsey T., Dear S., Devlin K., Fraser A.,
RA Gentles S., Hamlyn N., Horsnell T.S., Hunt S., Jagels K., Jones M.,
RA Louis E., Lye G., Moule S., Moule T., Odell C., Pearson D.,
RA Rajandream M.A., Riles L., Rowley N., Skelton J., Smith V.,
RA Walsh S.V., Whitehead S.;
RA Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Z46921; CAA87023.1; -.
CC SGD; S0001431; YIL169C.
CC InterPro; IPR004089; Chmtaxis_transd.
CC InterPro; IPR000727; T_SNARE.
CC PROSITE; PS50192; T_SNARE; UNKNOWN_1.
CC Hypothetical protein; Signal.
KW SIGNAL 1 23 POTENTIAL.
FT CHAIN 24 995 HYPOTHETICAL PROTEIN YIL169C.
FT DOMAIN 92 154 T-SNARE COILED-COIL HOMOMOLOGY (POTENTIAL).
FT CARBOHYD 28 28 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 35 35 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 468 468 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 664 664 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 995 AA; 99735 MW; F63E287A03F137EC CRC64;

Query Match 6.4%; Score 111; DB 1; Length 995;
```


Best Local Similarity 20.4%; Pred. No. 13;
Matches 68; Conservative 50; Mismatches 109; Indels 106; Gaps 17;
QY 4 ADAPAACTDIT---NQLTNVTVGIDS-----GTTVPYHQAGYVKLNFGFSPVNS 49
Db 222 SQASSTSDVSSSVSQASSTSGVSSSGSQSVSSASGSSSPQSTSSASTASGSAFNS 281
QY 50 AVKGDTEFKITVPKELNLNGVTSTAKVPPIMAGDQVLANGVIDSDGNVIYTFDYYVNTKDD 109
Db 282 -----LSSITSSAS-----SASATASNSLSSSDGTI-----YLPT-TT 313
QY 110 VKATLTPAYIDPENVKKTGNVTLATGIGSTTANKTVLVDEYKYGKFNLSIKGTIDQID 169
Db 314 ISGDLTL-----TGKVIATEGVVAAGAKLTLLDGDKYSFSADLKVYGD-LVK 361
QY 170 KTNNTYRQTIYVNPSSGDNVIAPVLTGNLKPNTDSNALIDQOQNTSIKVV-----KVDNAAD 224
Db 362 KSKETYPGTEF-DISGENF---DVTGNF--NAEESAA-----TSASIYFTPSSFDNSGD 410
QY 225 LSESYFVNPE---NFEDVTNSVNITFPNPNQYKVEFNTPDQITTPYIVVNGHIDPNK 281
Db 411 ISLSLSKSKKGEVTFSPYSNSGAFSPN-----AILNG-----GSV 446
QY 282 GDALRLSTLYGYNISNIIWRMSMSWDNEVAFNNGS 314
Db 447 SGLQRDDTEGSVNN-----GEINLDNGS 470

RESULT 19
SLAP_BACSH STANDARD; PRT; 1176 AA.
ID SLAP_BACSH
AC P38537;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE Surface-layer 125 kDa protein precursor.
OS Bacillus sphaericus.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1421;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=2362;
RX MEDLINE=89327128; PubMed=2666389;
RA Bowditch R.D., Baumann P., Yousten A.A.;
RT "Cloning and sequencing of the gene encoding a 125-kilodalton
RT surface-layer protein from Bacillus sphaericus 2362 and of a related
RT cryptic gene."
RL J. Bacteriol. 171:4178-4188(1989).
CC -!- FUNCTION: THE S-LAYER IS A PARACRYSTALLINE MONO-LAYERED ASSEMBLY
CC OF PROTEINS WHICH COAT THE SURFACE OF BACTERIA.
CC -!- SUBCELLULAR LOCATION: Cell wall.
CC -!- SIMILARITY: CONTAINS 3 S-LAYER HOMOMOLOGY (SLH) DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M28361; AAA50256.1; --
DR PIR; A33856; A33856.
DR InterPro; IPR001119; SLH.
DR Pfam; PF00395; SLH; 3.
DR PROSITE; PS01072; SLH_DOMAIN; 2.
KW Signal; Cell wall; S-layer; Repeat.
FT SIGNAL 1 30 POTENTIAL.
FT CHAIN 31 1176 SURFACE-LAYER 125 KDA PROTEIN.
FT DOMAIN 32 91 SLH 1.
FT DOMAIN 92 151 SLH 2.
FT DOMAIN 152 210 SLH 3.
SQ SEQUENCE 1176 AA; 125225 MW; 1A9A458EF433788C CRC64;

Query Match 6.4%; Score 110.5; DB 1; Length 1176;
Best Local Similarity 22.5%; Pred. No. 18;
Matches 72; Conservative 45; Mismatches 110; Indels 93; Gaps 16;
QY 28 TTVVPHQAGYVKLNFGFSPVNSAVKGDTEFKITVPKELNLNGVTSTAKVPPIMAGD--QVL 85
Db 157 STVKPWAQSYLEI---AVANGVIKGS-----EANGKTNLNPAPITRQDFAVVF 202
QY 86 ANGVIDS DG-----NVIYTFDYVNTKDDVKATLTMPAYIDPE-----N 124
Db 203 SRTIENVDPATPKVDKIEVVDKATLNVTLSDGTKETVLEKALEPNKETETVFKIKDVEYK 262
QY 125 VKKTGNVTLATGIGSTTAN--KTVLVDEYKYGKFNLSIKGTIDQ---IDKTNNTYRQTI 179
Db 263 AKVTYVVTATAVKSVSATNLKEVVVEFD-----GTVDKETAEDAA----- 304
QY 180 YVNPSSGDNVIAPVLTGNLKPNTDSNALIDQOQNTSIKVVKVDNAADLSESYFVNPFEDV 239
Db 305 YALKSGKTIKSVSLAADNK--TATVTLTDKLN-----NNKADA-----ISISNVKAG 349
QY 240 TNSVNITFPNPNQYKVEFNTPDQITTPYIVVNGHIDPNKSGDLALRLSTLYGYNISNIIW 299
Db 350 DKEINVK-----NVEFTAVDNKI--PEVTEV-----KSLGTKAVKVTLSPEVNLS 394
QY 300 RMSWDNEVAFNN---GSGS 316
Db 395 TNFTLDGKAYFGNVVMGAGN 414

RESULT 20
OMP_B_RICTY STANDARD; PRT; 1645 AA.
ID OMP_B_RICTY
AC P96989;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Outer membrane protein B precursor (168 kDa surface-layer protein)
DE (Surface protein antigen) (Cell surface antigen 5) (Sca5) (rOmpB)
DE (rOmp B) [Contains: 120 kDa surface-exposed protein (Surface protein
DE antigen) (120 kDa outer membrane protein ompB); 32 kDa beta peptide].
GN OMPB OR SLP.
OS Rickettsia typhi.
OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
OC Rickettsiaceae; Rickettsia.
OX NCBI_TaxID=785;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Wilmingon;
RX MEDLINE=94040787; PubMed=8224886;
RA Hahn M.-J., Kim K.-K., Kim I., Chang W.-H.;
RT "Cloning and sequence analysis of the gene encoding the crystalline
RT surface layer protein of Rickettsia typhi."
RL Gene 133:129-133(1993).
RN [2]
RP PARTIAL SEQUENCE.
RC STRAIN=Wilmingon;
RX MEDLINE=92114896; PubMed=1370573;
RA Ching W.M., Carl M., Dasch G.A.;
RT "Mapping of monoclonal antibody binding sites on CNR fragments of
RT the S-layer protein antigens of Rickettsia typhi and Rickettsia
RT prowazekii."
RL Mol. Immunol. 29:95-105(1992).
RN [3]
RP IDENTIFICATION OF CLEAVAGE SITE.
RX MEDLINE=92104668; PubMed=1729180;
RA Hackstadt T., Messer R., Cieplak W., Peacock M.G.;
RT "Evidence for proteolytic cleavage of the 120-kilodalton outer
RT membrane protein of rickettsiae: identification of an avirulent
RT mutant deficient in processing."
RL Infect. Immun. 60:159-165(1992).
CC -!- FUNCTION: THE 120 kDa SURFACE-EXPOSED PROTEIN IS A MAJOR
CC STRUCTURAL PROTEIN WHICH MAY PLAY A ROLE AS A RICKETTSIAL

QY 161 IKGTTIDQDKTNTYRTQTIYVNPSPGDNVIAPIVLTGNLKPNTDSNALIDQQNTSIKV---- 216

Db 179 -----DMLPDTHTVRWFLNINNEKSVYSKDI---TIKDQIQGGQQDLSTLNINVTGTH 230

QY 217 ---YKVDNA-ADLSESYFVNPFEDVNTSVNITFP-----NPNQYKVEFNT 259

Db 231 SNYYSQSAITDFEKAFFPGSKITVDNTKNTIDVTIPQGYGSYNSFSINYKT 281

RESULT 22

SIX4_HUMAN

AC Q9UIU6; STANDARD; PRT; 760 AA.

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Homeobox protein SIX4 (Sine oculis homeobox homolog 4).

GN SIX4.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=20108802; PubMed=10640827;

RA Ozaki H., Yamada K., Kobayashi M., Asakawa S., Minoshima S., Shimizu N., Kajitani M., Kawakami K.;

RT "Structure and chromosome mapping of the human SIX4 and murine Six4 genes.";

RL Cytogenet. Cell Genet. 87:108-112(1999).

CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).

CC -!- SIMILARITY: BELONGS TO THE SIX/SINE OCULIS HOMEOBOX FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/) or send an email to license@isb-sib.ch).

CC -----

DR EMBL; AB024687; BAA86223.1; -.

DR EMBL; AB024685; BAA86223.1; JOINED.

DR EMBL; AB024686; BAA86223.1; JOINED.

DR HSSP; P41778; 1DU6.

DR TRANSFAC; T03344; -.

DR Genew; HGNC:10890; SIX4.

DR MIM; 606342; -.

DR InterPro; IPR001356; Homeobox.

DR Pfam; PF00046; homeobox; 1.

DR ProDom; PD000010; Homeobox; 1.

DR SMART; SM00389; HOX; 1.

DR PROSITE; PS00027; HOMEBOX_1; 1.

DR PROSITE; PS50071; HOMEBOX_2; 1.

KW Developmental protein; Homeobox; DNA-binding; Nuclear protein.

FT DNA_BIND 202 261 HOMEBOX.

SQ SEQUENCE 760 AA; 80832 MW; 9B0CADD17498D30 CRC64;

Query Match 6.3%; Score 109.5; DB 1; Length 760;

Best Local Similarity 21.6%; Pred. No. 11;

Matches 81; Conservative 45; Mismatches 114; Indels 135; Gaps 19;

- QY 30 VYPHAGYVKLNY-----GFSVPNSA-----VKGDTFKITVPKELNLNGVT-----ST 72

Db 310 VYMQQIGNAKISLSSSGVLLNGSLVPASTSPVFLNGNSF-IQGPGSVILNGLNVGNTQAV 368

QY 73 AKVPPIMAGDQVLANGVIDSGNVIYTFDYV-NTKDDVK-----ATLT----- 115

Db. 369 ALNPPKMSSN-IVSNGI-----SMTDILGSTQDVKEFKVLQSSANSATTTSYSPS 418

QY 116 -----MPAYIDPENVKKTGNVTPLATGIGSTTANKTTLVDYKEYG-----KPYNLS 160

Db 419 VPVSFPGLIPSTEVKREGIQTVASQDGGSVVTFTPVQINOYGIQIPN\$GANSQFLNGS 478

QY 161 I-----KGTIDQIDKTNN-----TYRQ-----TIYVNPS 184

Db 479 IGFSPQLQPPVSVAA\$QGNISVSSTSDGSTFTSESTTVQ\$QGVFLSSLAPSAVVYTPVN 538

QY 185 GDNVIAPIVLTGNLK-----PNTDSNALIDQQNTS-----IKVYKVDN 221

Db 539 TGQTIGSVKQEGLESLVFSQLMPVNQNAQVNNLSSNISGLPPLASSLVNVSPTHN 598

QY 222 AADLSESYFVNP-ENFEDVNTSVNITFP-----NPNQYKVE---FNTPPDDQITTT 266

Db 599 FS-LSPSTLLNPTLNRDIAD\$QPM\$APVASKSTVTSVNTNYATLQNC\$SLITQ\$D\$LLSV 657

QY 267 PYIVVNVNGHIDPNSK 281

Db 658 PMTQAALGEIVPTAE 672

RESULT 23

YDBA_ECOLI

ID YDBA_ECOLI STANDARD; PRT; 2003 AA.

AC P33666; P76087; P76088; P76856; P76857; P76859;

DT 01-FEB-1994 (Rel. 28, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Hypothetical protein ydba.

GN YDBA OR B1401/B1405.

OS Escherichia coli.

OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; Escherichia.

OX NCBI_TaxID=562;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=K12 / MCL655;

RX MEDLINE=97426617; PubMed=9278503;

RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V., Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F., Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J., Mau B., Shao Y.;

RT "The complete genome sequence of Escherichia coli K-12.";

RL Science 277:1453-1474(1997).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=K12;

RX MEDLINE=97251357; PubMed=9097039;

RA Aiba H., Baba T., Fujita K., Hayashi K., Inada T., Isono K., Itoh T., Kasai H., Kashimoto K., Kimura S., Kitakawa M., Kitagawa M., Makino K., Miki T., Mizobuchi K., Mori H., Mori T., Motomura K., Nakade S., Nakamura Y., Nashimoto H., Nishio Y., Oshima T., Saito N., Sampei G., Seki Y., Sivasubram S., Tagami H., Takeda J., Takemoto K., Takeuchi Y., Wada C., Yamamoto Y., Horiuchi T.;

RT "A 570-kb DNA sequence of the Escherichia coli K-12 genome corresponding to the 28.0-40.1 min region on the linkage map.";

RL DNA Res. 3:363-377(1996).

RN [3]

RP SEQUENCE OF 464-2003 FROM N.A.

RC STRAIN=K12;

RX MEDLINE=92190338; PubMed=1665988;

RA Moszer I., Glaser P., Danchin A.;

RT "Multiple IS insertion sequences near the replication terminus in Escherichia coli K-12.";

RL Biochimie 73:1361-1374(1991).

CC -!- SIMILARITY: TO S.TYPHIMURIUM ORF NEAR CYSG (AC P25928).

CC -!- CAUTION: THIS IS A CONCEPTUAL TRANSLATION; THE GENE CODING FOR THIS PROTEIN IS INTERRUPTED BY A HYBRID IS2D/IS30C ELEMENT BETWEEN AMINO ACIDS 839 AND 840.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 23, 2003, 13:59:10 ; Search time 42.8422 Seconds
(without alignments)
1591.926 Million cell updates/sec

Title: US-10-056-052A-4
Perfect score: 1732
Sequence: 1 MVAADAPAAAGTDITNQLTNV.....NGSGSGDIDKPVVPEQPDE 331

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues 671580
1 number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

- Database : SPTREMBL_21:*
- 1: sp_archaea:*
 - 2: sp_bacteria:*
 - 3: sp_fungi:*
 - 4: sp_human:*
 - 5: sp_invertebrate:*
 - 6: sp_mammal:*
 - 7: sp_mhc:*
 - 8: sp_organelle:*
 - 9: sp_phage:*
 - 10: sp_plant:*
 - 11: sp_rodent:*
 - 12: sp_virus:*
 - 13: sp_vertebrate:*
 - 14: sp_unclassified:*
 - 15: sp_rvirus:*
 - 16: sp_bacteriaph:*
 - 17: sp_archaeap:*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1727	99.7	933	2	Q53653
2	1517	87.6	935	16	Q932C5
3	1517	87.6	989	16	Q99VJ4
4	1112	64.2	881	2	Q93MH7
5	416.5	24.0	961	16	Q99RD3
6	409	23.6	940	2	Q53682
7	353	20.4	931	2	Q9K1I3
8	349	20.2	1092	2	O70022
9	324.5	18.7	1038	16	Q99RD2
10	323.5	18.7	1166	2	O86489
11	320.5	18.5	1141	16	Q99W46
12	320.5	18.5	1141	16	Q932F7
13	312.5	18.0	877	16	Q99R07
14	312.5	18.0	913	2	O86476
15	298	17.2	1171	2	Q9KWX6
16	257.5	14.9	1315	2	O86488

90	116	6.7	784	16	Q8YAJ5	Q8yaj5 listeria mo
91	116	6.7	839	2	P77792	P77792 chlamydia p
92	116	6.7	1570	5	Q9U0H8	Q9u0h8 plasmodium
93	116	6.7	2340	5	O97298	O97298 plasmodium
94	115.5	6.7	461	17	Q8TTC5	Q8ttc5 methanosarc
95	115.5	6.7	1603	2	Q9KKA6	Q9kka6 rickettsia
96	115.5	6.7	2806	16	Q8RI19	Q8ril9 fusobacteri
97	115.5	6.7	4936	16	Q8YKJ3	Q8ykJ3 anabaena sp
98	115	6.6	785	16	Q92FC9	Q92fc9 listeria in
99	115	6.6	858	5	P90545	P90545 entamoeba h
100	115	6.6	1582	16	Q8Y9A5	Q8y9a5 listeria mo

ALIGNMENTS

```
RESULT 1
Q53653
ID Q53653 PRELIMINARY; PRT; 933 AA.
Q53653;
01-NOV-1996 (TrEMBLrel. 01, Created)
01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Clumping factor.
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Staphylococcus.
OX NCBI_TaxID=1280;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NEWMAN;
RX MEDLINE=94224142; PubMed=8170386;
RA McDevitt D., Francois P., Vaudaux P., Foster T.J.;
RT "Molecular characterization of the clumping factor(fibrogen receptor
of Staphylococcus aureus.";
RL Mol. Microbiol. 11:237-248(1994).
DR EMBL; Z18852; CAA79304.1; -.
DR InterPro; IPR000515; BPD_transp.
DR InterPro; IPR001899; Gram_pos_anchor.
DR TIGRFAMS; TIGR01167; LPXTG_anchor; 1.
DR PROSITE; PS00402; BPD_TRANS_INN_MEMBR; UNKNOWN_1.
DR PROSITE; PS00343; GRAM_POS_ANCHORING; UNKNOWN_1.
SQ SEQUENCE 933 AA; 97058 MW; EB51A6DE2FF759F4 CRC64;
```

```
Query Match 99.7%; Score 1727; DB 2; Length 933;
Best Local Similarity 100.0%; Pred. No. 1.9e-87;
Matches 330; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Qy	2	VAADAPAAAGTDITNQLTNVTVGIDSGTTVYPHQAGYVKLYGFSVPNSAVKGDTEKITVP	61
Db	221	VAADAPAAAGTDITNQLTNVTVGIDSGTTVYPHQAGYVKLYGFSVPNSAVKGDTEKITVP	280
Qy	62	KELNLNGVTSTAKVPPIMAGDQVLANGVIDSDGNVIYFTFDYVNTKDDVKATLTMPAYID	121
Db	281	KELNLNGVTSTAKVPPIMAGDQVLANGVIDSDGNVIYFTFDYVNTKDDVKATLTMPAYID	340
Qy	122	PENVKKTGNVTLATGIGSTTANKTVLVDYKYGKFNLSIKGTIDQIDKTNNTYRQTIYV	181
Db	341	PENVKKTGNVTLATGIGSTTANKTVLVDYKYGKFNLSIKGTIDQIDKTNNTYRQTIYV	400
Qy	182	NPSGDNVIAPVLTGNLKPNTDSNALIDQONTSIKVKYKVDNAADLSESYFVNPFEDVTN	241
Db	401	NPSGDNVIAPVLTGNLKPNTDSNALIDQONTSIKVKYKVDNAADLSESYFVNPFEDVTN	460
Qy	242	SVNITFPNPNOYKVEFNTPDQITTPYIVVNGHIDPNKSGDLALRSTLYGYSNIIWRS	301
Db	461	SVNITFPNPNOYKVEFNTPDQITTPYIVVNGHIDPNKSGDLALRSTLYGYSNIIWRS	520
Qy	302	MSWDNEVAFNNGSGSGDGIDKPVVPEQPDE	331
Db	521	MSWDNEVAFNNGSGSGDGIDKPVVPEQPDE	550

```
RESULT 2
Q932C5
ID Q932C5 PRELIMINARY; PRT; 935 AA.
AC Q932C5;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Fibrinogen-binding protein.
GN FNB OR SAV0811.
OS Staphylococcus aureus (strain Mu50 / ATCC 700699).
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Staphylococcus.
OX NCBI_TaxID=158878;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21311952; PubMed=11418146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,
RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,
RA Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
RA Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,
RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;
RT "Whole genome sequencing of methicillin-resistant Staphylococcus
aureus.";
RL Lancet 357:1225-1240(2001).
DR EMBL; AP003360; BAB56973.1; -.
DR InterPro; IPR000515; BPD_transp.
DR InterPro; IPR001899; Gram_pos_anchor.
DR TIGRFAMS; TIGR01167; LPXTG_anchor; 1.
DR TIGRFAMS; TIGR01168; YSIRK_signal; 1.
DR PROSITE; PS00402; BPD_TRANS_INN_MEMBR; UNKNOWN_1.
DR PROSITE; PS00343; GRAM_POS_ANCHORING; UNKNOWN_1.
KW Complete proteome.
SQ SEQUENCE 935 AA; 96950 MW; DC5A2D92CE3BA91C CRC64;
```

```
Query Match 87.6%; Score 1517; DB 16; Length 935;
Best Local Similarity 87.3%; Pred. No. 7.3e-76;
Matches 288; Conservative 17; Mismatches 25; Indels 0; Gaps 0;
```

Qy	2	VAADAPAAAGTDITNQLTNVTVGIDSGTTVYPHQAGYVKLYGFSVPNSAVKGDTEKITVP	61
Db	221	VAADAPAAAGTDITNQLTDVKVTIDSGTTVYPHQAGYVKLYGFSVPNSAVKGDTEKITVP	280
Qy	62	KELNLNGVTSTAKVPPIMAGDQVLANGVIDSDGNVIYFTFDYVNTKDDVKATLTMPAYID	121
Db	281	KELNLNGVTSTAKVPPIMAGDQVLANGVIDSDGNVIYFTFDYVNDKENVTANITMPAYID	340
Qy	122	PENVKKTGNVTLATGIGSTTANKTVLVDYKYGKFNLSIKGTIDQIDKTNNTYRQTIYV	181
Db	341	PENVTKTGNVTLTTGIGTNTASKTVLIDYKYGQGFHNLSIKGTIDQIDKTNNTYRQTIYV	400
Qy	182	NPSGDNVIAPVLTGNLKPNTDSNALIDQONTSIKVKYKVDNAADLSESYFVNPFEDVTN	241
Db	401	NPSGDNVVLPAITGNLIPNPKSNALIDAKNTDIKVYRVDNANDLSESYVNPSPDEFVTN	460
Qy	242	SVNITFPNPNOYKVEFNTPDQITTPYIVVNGHIDPNKSGDLALRSTLYGYSNIIWRS	301
Db	461	QVRISFPNANQYKVEFPTDDQITTPYIVVNGHIDPASTGDLALRSTFYGYDSNFIWRS	520
Qy	302	MSWDNEVAFNNGSGSGDGIDKPVVPEQPDE	331
Db	521	MSWDNEVAFNNGSGSGDGIDKPVVPEQPDE	550

```
RESULT 3
Q99VJ4
ID Q99VJ4 PRELIMINARY; PRT; 989 AA.
AC Q99VJ4;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Fibrinogen-binding protein A, clumping factor.
```

```
GN CLFA OR SA0742.
OS Staphylococcus aureus (strain N315).
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Staphylococcus.
OX NCBI_TaxID=158879;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21311952; PubMed=11418146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J., Ito T., Kanamori M.,
RA Matsumaru H., Maruyama A., Murakami H., Hosoyama A., Mizutani-Ui Y.,
RA Takahashi N.K., Sawano T., Inoue R.I., Kaito C., Sekimizu K.,
RA Hirakawa H., Kuhara S., Goto S., Yabuzaki J., Kanehisa M.,
RA Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T., Hattori M.,
RA Ogasawara N., Hayashi H., Hiramatsu K.;
RT "Whole genome sequencing of methicillin-resistant staphylococcus
RT aureus.";
RL Lancet 357:1225-1240(2001).
DR EMBL; AP003131; BAB41975.1; -.
InterPro; IPR000515; BPD_transp.
InterPro; IPR001899; Gram_pos_anchor.
TIGRFAMS; TIGR01167; LPXTG_anchor; 1.
PROSITE; PS00402; BPD_TRANSF_INN_MEMBER; UNKNOWN_1.
DR PROSITE; PS00343; GRAM_POS_ANCHORING; UNKNOWN_1.
KW Complete proteome.
SQ SEQUENCE 989 AA; 102407 MW; DA6E807539623467 CRC64;

Query Match 87.6%; Score 1517; DB 16; Length 989;
Best Local Similarity 87.3%; Pred. No. 7.8e-76;
Matches 288; Conservative 17; Mismatches 25; Indels 0; Gaps 0;

QY 2 VAADAPAAAGTDITNQLTNVTVGIDSGTTVYPHQAGYVKLYGFSVPNSAVKGDTEKITVP 61
Db 221 VAADAPAAAGTDITNQLTDVKVTIDSGTTVYPHQAGYVKLYGFSVPNSAVKGDTEKITVP 280
QY 62 KELNLNGVTSTAKVPPIMAGDQVLANGVIDSDGNVIYFTFDYVNTKDDVKATLTMPAYID 121
Db 281 KELNLNGVTSTAKVPPIMAGDQVLANGVIDSDGNVIYFTFDYVDNKENVTANITMPAYID 340
QY 122 PENVKKTGNVTLATGIGSTTANKTVLVVDYKEYGKFYNLSIKGTIDQIDKTNNTYRQTIYV 181
Db 341 PENVTKTGNVTLTGIGTNTASKTVLIDYKEYGQFHNLSIKGTIDQIDKTNNTYRQTIYV 400
QY 182 NPSGDNVIAPVLTGNLKPNTDSNALIDQNTSIKVKYKVDNAADLSESYFVNPFEDVTN 241
Db 401 NPSGDNVVLPAITGNLIPNTKSNALIDAKNTDIKVRVDNANDLSESYFVNPSDFEDVTN 460
QY 242 SVNITFPNPQYKVEFNTPDQITTPYIVVNGHIDPNKSGDLALRSTLYGYNINSNIWRS 301
Db 461 QVRISFPNANQYKVEFPTDDQITTPYIVVNGHIDPASTGDLALRSTFYGYDSNFIWRS 520
QY 302 MSWDNEVAFNNGSGSGDGIDKPVVPEQPDE 331
Db 521 MSWDNEVAFNNGSGSGDGIDKPVVPEQPDE 550

RESULT 4
Q93MH7 PRELIMINARY; PRT; 881 AA.
AC Q93MH7;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Receptin Fbl precursor.
GN FBL.
OS Staphylococcus lugdunensis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Staphylococcus.
OX NCBI_TaxID=28035;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=2342;
RA Nilsson M., Ahlen J., Frykberg L., Guss B.;
```

```
RT "A fibrinogen-binding protein of Staphylococcus lugdunensis.";
RL Submitted (AUG-2001) to the EMBL/GenBank/DBDJ databases.
DR EMBL; AF404823; AAK95649.1; -.
DR InterPro; IPR001899; Gram_pos_anchor.
DR Pfam; PF00746; Gram_pos_anchor; 1.
DR TIGRFAMS; TIGR01167; LPXTG_anchor; 1.
DR TIGRFAMS; TIGR01168; YSIRK_signal; 1.
DR PROSITE; PS00343; GRAM_POS_ANCHORING; UNKNOWN_1.
KW Signal.
FT SIGNAL 1 39 POTENTIAL.
FT CHAIN 40 881 RECEPTIN FBL.
SQ SEQUENCE 881 AA; 94251 MW; D4296C4959C4F19B CRC64;

Query Match 64.2%; Score 1112; DB 2; Length 881;
Best Local Similarity 62.6%; Pred. No. 1.5e-53;
Matches 201; Conservative 50; Mismatches 70; Indels 0; Gaps 0;

QY 11 TDITNQLTNVTVGIDSGTTVYPHQAGYVKLYGFSVPNSAVKGDTEKITVPKELNLNGVT 70
Db 214 SDISNKLNSVTATIEAADTIYPHKAEYVNLNRYFQAPDDVQAGDSIKITIPQALNLNGVT 273
QY 71 STAKVPPIMAGDQVLANGVIDSDGNVIYFTFDYVNTKDDVKATLTMPAYIDPENVKKTGN 130
Db 274 ATAKAPNIMAGDQILATGTIDEENLIYFTFDYVNTKNNITGOISIPGVIDPKNVTHTGK 333
QY 131 VTLATGIGSTTANKTVLVVDYKEYGKFYNLSIKGTIDQIDKTNNTYRQTIYVNPSSDGNVIA 190
Db 334 VNLETSIGQTTAKKTVVDYKEYGFEFRNLISIKGTIDQIDKVNNTYRQTIYVNPSSDVTVD 393
QY 191 PVLGNLKPNTDSNALIDQNTSIKVKYKVDNAADLSESYFVNPFEDVTNSVNITFPNP 250
Db 394 PYLRGGSIPGNTSNVVIDEQNTSIKVKYKVEKAHLTDSYVVDPSNYEDVTSVKITFPEK 453
QY 251 NQYKVEFNTPDQITTPYIVVNGHIDPNKSGDLALRSTLYGYNINSNIWRSMSWDNEVAF 310
Db 454 GIYQINFENTEDDQINSYVNVINGHVDPNNGNLRLSTLYGYSNFTRVSMADNEVEY 513
QY 311 NNGSGSGDGIDKPVVPEQPDE 331
Db 514 HAGSGNGDGIDKPVIPDQPGD 534

RESULT 5
Q99RD3 PRELIMINARY; PRT; 961 AA.
AC Q99RD3;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE FnbB protein (Fibroectin-binding protein homolog).
GN FNB OR SAV2502 OR SA2290.
OS Staphylococcus aureus (strain Mu50 / ATCC 700699), and
OC Staphylococcus aureus (strain N315).
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Staphylococcus.
OX NCBI_TaxID=158878, 158879;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=S.aureus (strain Mu50), and S.aureus (strain N315);
RX MEDLINE=21311952; PubMed=11418146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,
RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,
RA Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
RA Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,
RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;
RT "Whole genome sequencing of methicillin-resistant Staphylococcus
RT aureus.";
RL Lancet 357:1225-1240(2001).
DR EMBL; AP003365; BAB58664.1; -.
DR EMBL; AP003137; BAB43593.1; -.
DR InterPro; IPR004237; Fn_bind.
```


QY 187 NVIAPVLGTGNLKPNTDSNALIDQQNTSIKVKYKVDNAADLSESYFV-NPENFEDVTNSVNI 245
Db 459 RYSAKETNVNISGNGDEGSTIIDSTIIKVKYKVDNQNLPSNRIYDYSEYEDVTNDYYA 518
.QY 246 TFPNPNQYKVEFNTPDQITTPYIVVNVNGHIDPNSKGD-----LALRSTLYGNSNII 298
Db 519 QLGNNNDVNINEG---NIDSPYIIKVISKYDPN-KDDYTTIQQVTVMQTTINEYTG-- 571
.QY 299 WRSMSWDNEVAFNNGSGSGDGIDKPKVPVPEQ 328
Db 572 FRTASYDNTIAFSTSSGGQG-DLP--PEK 598

* RESULT 8
O70022 PRELIMINARY; PRT; 1092 AA.
ID O70022
AC O70022;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Fibrinogen-binding protein precursor.
OS Staphylococcus epidermidis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Staphylococcus.
OX NCBI_TaxID=1282;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HB;
RX MEDLINE=98261511; PubMed=9596732;
RA Nilsson M., Frykberg L., Flock J.I., Pei L., Lindberg M., Guss B.;
RT "A Fibrinogen-binding protein of Staphylococcus epidermidis.";
RL Infect. Immun. 66:2666-2673(1998).
DR EMBL; Y17116; CAA76638.1; -;
DR InterPro; IPR001899; Gram_pos_anchor.
DR Pfam; PF00746; Gram_pos_anchor; 1.
DR TIGRFAMS; TIGR01167; LPXTG_anchor; 1.
DR TIGRFAMS; TIGR01168; YSIRK_signal; 1.
DR PROSITE; PS00343; GRAM_POS_ANCHORING; UNKNOWN_1.
KW Signal.
FT SIGNAL 1 51 POTENTIAL.
FT CHAIN 52 1092 FIBRINOGEN-BINDING PROTEIN.
SQ SEQUENCE 1092 AA; 119292 MW; 6542BC39AAD8B984 CRC64;

Query Match 20.2%; Score 349; DB 2; Length 1092;
Best Local Similarity 30.3%; Pred. No. 2.3e-11;
Matches 100; Conservative 61; Mismatches 139; Indels 30; Gaps 14;

13 ITNQLTNVTGI-DSGTTVYPHQAGYVKLYGFSVPNSAVKGDTEKITVPKELNLNGVTS 71
287 VTQD--SITEGDDSEGVKAHDAENLIYDVFEDDKVKSGDTMTVDIDKNTVPSDLTD 344

72 TAKVPPIMAGD-QVLANGVIDS-DGNVIYTFDYVNTKDDVKATLTMPAYIDPENVKKTG 129
345 SFTIPKIKDNGSGEIIATGTYDNKNKQIYTFDYVDKYENIKAHLKLTSYIDKSKVPNNN 404

130 ---NVTLATGIGSTTANKTVLDYEKYGFYNLSIKGTIDQIDKTNNTYRQTIYVNPSPGD 186
405 TKLDVEYKTALSS--VNKTIITVEYQRPNENRTANLQSMFTNIDTKNHTVEQTIYINPL-- 460

187 NVIAPVLGTGNLKPNTDSNALIDQQNTSIKVKYKVDNAADLSESYFV-NPENFEDVTNSVNI 245
461 RYSAKETNVNISGNGDEGSTIIDSTIIKVKYKVDNQNLPSNRIYDYSEYEDVTNDYYA 520
246 TFPNPNQYKVEFNTPDQITTPYIVVNVNGHIDPNSKGD-----LALRSTLYGNSNII 298
521 QLGNNNDVNINEG---NIDSPYIIKVISKYDPN-KDDYTTIQQVTVMQTTINEYTG-- 573
299 WRSMSWDNEVAFNNGSGSGDGIDKPKVPVPEQ 328
574 FRTASYDNTIAFSTSSGGQG-DLP--PEK 600

RESULT 9
Q99RD2 PRELIMINARY; PRT; 1038 AA.
AC Q99RD2;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Fnb protein (Fibronectin-binding protein homolog).
GN FNB OR SAV2503 OR SA2291.
OS Staphylococcus aureus (strain Mu50 / ATCC 700699), and
OS Staphylococcus aureus (strain N315).
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Staphylococcus.
OX NCBI_TaxID=158878, 158879;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=S.aureus (strain Mu50), and S.aureus (strain N315);
RX MEDLINE=21311952; PubMed=11418146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,
RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,
RA Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
RA Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,
RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;
RT "Whole genome sequencing of meticillin-resistant Staphylococcus
RT aureus.";
RL Lancet 357:1225-1240(2001).
DR EMBL; AP003365; BAB58665.1; -;
DR EMBL; AP003137; BAB43594.1; -;
DR InterPro; IPR004237; Fb_bind.
DR InterPro; IPR001899; Gram_pos_anchor.
DR InterPro; IPR002965; P_rich_extensn.
DR Pfam; PF02986; Fb_bind; 1.
DR Pfam; PF00746; Gram_pos_anchor; 1.
DR PRINTS; PR01217; PRICHEXTENS.
DR TIGRFAMS; TIGR01167; LPXTG_anchor; 1.
DR TIGRFAMS; TIGR01168; YSIRK_signal; 1.
DR PROSITE; PS00343; GRAM_POS_ANCHORING; UNKNOWN_1.
KW Complete proteome.
SQ SEQUENCE 1038 AA; 113618 MW; 666BF6BF2BFBE12 CRC64;

Query Match 18.7%; Score 324.5; DB 16; Length 1038;
Best Local Similarity 26.9%; Pred. No. 4.8e-10;
Matches 91; Conservative 68; Mismatches 154; Indels 25; Gaps 11;

QY 10 GTDITNQLTNVTVGIDS--GTTVYPHQAGYVKLYGFSVPNSAVKGDTEKITVPKELNLN 67
Db 189 GTDVTSKVTVESGSEIAPQGNKVEPHAGQVRVLKYKLFADGLKRGDYDFDTLSNNVNTY 248

QY 68 GVTSTAKVPPIMAGDQVLANGVIDSDGNVIYTFDYVNTKDDVKATLTMPAYIDPENVKK 127
Db 249 GVSTARKVPEIKNGSVVMATGEILGNGNIRYTFETNEIEHKVEVTANLEINFIDPKTVQS 308

QY 128 TGNVTLATGIGSTTANKTVLDYEKYGFYNLSIKGTIDQIDKTNNTYRQTIYVNPSPGD 187
Db 309 NGEQKITSKLNGETEKTIPVVYNPGVSNSTYTNVNGSIETFNKESNKFTTHIAYIKPMNGN 368

QY 188 VIAPV-LTGNLKPNTDSNALIDQQNTSIKVKY-VDNAADLSESYFVNPN---FEDVTNS. 242
Db 369 QSNTVSVTGTL---TEGSNLAGGQPT-VKVVEYLGKKDELQPSVYANTSDTNKFKDVTKE 424

QY 243 VN--ITFPNPNQYKVEFNTPDQITTPYIVVNVNGHIDPNSKGDALRSTLYG-----YNS 295
Db 425 MNGKLSVQDNGSGSYSLNL---DKLDKTYVIHYTGEYLOGS-DQVNFTELYGYPERAYKS 479

QY 296 NIIWRS--MSWDNEVAFNNGSGSGDGIDKPKVPVPEQ 331
Db 480 YVYGGYRLTWDNGLVLYSNKADGNKGNGQIIQDNDFE 517

RESULT 10
O86489

RT "Whole genome sequencing of meticillin-resistant Staphylococcus
aureus."
RL Lancet 357:1225-1240(2001).
DR EMBL; AP003359; BAB56725.1; -.
DR InterPro; IPR001899; Gram_pos_anchor.
DR Pfam; PF00746; Gram_pos_anchor; 1.
DR TIGRFAMS; TIGR01167; LPXTG_anchor; 1.
DR TIGRFAMS; TIGR01168; YSIRK_signal; 1.
DR PROSITE; PS00343; GRAM_POS_ANCHORING; UNKNOWN_1.
KW Complete proteome.
SQ SEQUENCE 1141 AA; 124038 MW; E679F7C2991846D9 CRC64;

Query Match 18.5%; Score 320.5; DB 16; Length 1141;
Best Local Similarity 30.4%; Pred. No. 9e-10;
Matches 108; Conservative 55; Mismatches 139; Indels 53; Gaps 17;

QY 2 VAADAPAAAGTDITNQLT---NVTVGIDSGTTVYPHQAGYVKLNLYGFSVPNSAVKGDTEK 57
|| | | : : : : : || | : : : : : ||||
Db 264 VAQPAAVASNNVNDLITVTQTIKVGDKGNVAAAHDKGDI EYDTEFTIDNKKVKKGMT 323
|| | | : : : : : || | : : : : : ||||
58 I-----TVPKELNLNGVTSKAVPPIMAGD---QVLANGVID-SDGNVIYFTFDYVNTKD 108
| | | | : | | | : | | | : | | | : | | | | :
Db 324 INVDNKVI PSDL-----TDKNPDIDITDPSGEVIAKGFDFKATKQIYTFYTFDYVDKYE 376
| | | | : | | | : | | | : | | | : | | | :
QY 109 DVKATLTMPAYIDPENV--KKTGNVTLATGIGSTTANKTVLVVDYEKYKGFYNLSIKGTID 166
| | | | : | | | : | | | : | | | : | | | :
Db 377 DIKARLTLYSIDKQAVPNFETSLNLTATAGKETSQN--VSVDYQDPMVHGDSNIQSIFT 434
| | | | : | | | : | | | : | | | : | | | :
QY 167 QIDKTNNTYQTIYVNP---SGDNVIAPVL-----TGNLKPNTDSNALIDQNTSIKVV 217
| | | | : | | | : | | | : | | | : | | | |
Db 435 KLDENKQTI EQIYVNP LKKTATNTKVDIAGSQVDDYGNKLGNGS-TIID-QNTEIKVV 492
| | | | : | | | : | | | : | | | : | | | :
QY 218 KVDNAADLSESYFV-NPENFEDVNTSVNITFPNPNOYKVEFNTPD-DQITTPYIVVNGH 275
| | | | : | | | : | | | : | | | : | | | :
Db 493 KVPNQQLPQSNRIYDFSEYEDVTSQ-----FDNKKSFSSNNVATLDFGDINSAYIIKVVSK 548
| | | | : | | | : | | | : | | | : | | | :
QY 276 IDPNSKGDL-----ALRST-LYGYNSNIIWRMSWDNEVAFNNGSGSGDGIDKP 323
| | | | : | | | : | | | : | | | : | | | |
Db 549 YTPSDGELDIAQGTSMRTTDKYG-----YNYAGYSNFIVTSNDTGGGDTGTVKP 598
| | | | : | | | : | | | : | | | : | | | |

RESULT 13
Q99R07 PRELIMINARY; PRT; 877 AA.
ID Q99R07
AC Q99R07;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Clumping factor B.
GN CLFB OR SAV2630 OR SA2423.
OS Staphylococcus aureus (strain Mu50 / ATCC 700699), and
OS Staphylococcus aureus (strain N315).
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Staphylococcus.
OX NCBI_TaxID=158878, 158879;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=S.aureus (strain Mu50), and S.aureus (strain N315);
RX MEDLINE=21311952; PubMed=11418146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,
RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,
RA Mizutani-U I. Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
RA Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,
RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;
RT "Whole genome sequencing of meticillin-resistant Staphylococcus
aureus."
RL Lancet 357:1225-1240(2001).
DR EMBL; AP003366; BAB58792.1; -.
DR EMBL; AP003138; BAB43728.1; -.
DR InterPro; IPR001899; Gram_pos_anchor.
DR Pfam; PF00746; Gram_pos_anchor; 1.

DR TIGRFAMS; TIGR01167; LPXTG_anchor; 1.
DR TIGRFAMS; TIGR01168; YSIRK_signal; 1.
DR PROSITE; PS00343; GRAM_POS_ANCHORING; UNKNOWN_1.
KW Complete proteome.
SQ SEQUENCE 877 AA; 93649 MW; F0EF0A57AFDB357F CRC64;

Query Match 18.0%; Score 312.5; DB 16; Length 877;
Best Local Similarity 24.9%; Pred. No. 1.8e-09;
Matches 92; Conservative 69; Mismatches 132; Indels 77; Gaps 15;

QY 6 APAAGTDITNQLTNVTVGIDSGTTVYPHQAGYVKLNLYGFSVPNSAVKGDTEFTKITVPKELN 65
| | | | : : | | : : | | : : | | : : | | : : | |
Db 208 ADAKGTNVNDKVTASNFKLEK-TTFDPNQSGNTFMAANFTTVDKVKSGDYF----- 257
| | | | : : | | : : | | : : | | : : | | : : | |
QY 66 LNGVTSTAKVPPIMAGDQVLANGVID-----SDGNVI----- 97
| | | | : : | | : : | | : : | | : : | | : : | |
Db 258 -----TAKLP-----DSLTCNGDGDVYSNNTMTPIADIKSTNGDVVAKATYDILT KTYT 306
| | | | : : | | : : | | : : | | : : | | : : | |
QY 98 YTFDYVNTKDDVKATLTMPAYIDPENVKKTGNVTLATGIGSTTANKTVLVVDYEK-----Y 153
: | | | | | : : : : | | : : | | : : | | : : | |
Db 307 FVFTDYVNKENINGQFSLPFTDRAKAPKSGTYDANINIADEMFNKKITYNYSSPIAGI 366
| | | | : | | | : | | | : | | | : | | | : | | | :
QY 154 GKFNLSIKGTIDQIDKTN--NTYRQTIYVNP-----GDNVIAPVLTGNLKPNTDSNALI 207
| | | | : | | | : | | | : | | | : | | | : | | | :
Db 367 DKPNGANISSQIIGVDTASQNTYKQTVFVNPQRVLTGNTWV--YIKGYQDKIEESSGKV 424
| | | | : | | | : | | | : | | | : | | | : | | | :
QY 208 DQNTSIKVYKVDNAADLSESYFVNP--ENFEDVNTSV--NITFPNPNOYKVEFNTPDQD 263
: | | | | : : | | : : | | : : | | : : | | : : | |
Db 425 SATDTKLRIFEVNDTSKLSDSYVADPNDSNLKEVTDQFKNRIYEHHPNVA SIKFG----D 480
| | | | : | | | : | | | : | | | : | | | : | | | :
QY 264 ITTPYIVVNGHIDPNSKGLALRSTLYGYN----SNIIWRMSWDNEVAFNNGSGSGDG 319
| | | | : | | | : | | | : | | | : | | | : | | | :
Db 481 ITKTYVVLVEGHYDNTGKN--LKTQVIQENVDPVTNRDYSIFGWNNENVVRYGGGSADG 537
| | | | : | | | : | | | : | | | : | | | : | | | :
QY 320 IDKPVVPEQP 329
| | | | : | | | : | | | : | | | : | | | : | | | :
Db 538 -DSAVNPKDP 546
| | | | : | | | : | | | : | | | : | | | : | | | :

RESULT 14
O86476 PRELIMINARY; PRT; 913 AA.
ID O86476
AC O86476;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Clumping factor B precursor.
GN CLFB.
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Staphylococcus.
OX NCBI_TaxID=1280;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NEWMAN;
RA Foster T.J.;
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=NEWMAN;
RA N. Eidhin D., Perkins S., Francois P., Vaudaux P.;
RT "Clumping factor B(Clfb), a new surface-located fibrinogen-binding
adhesin of Staphylococcus aureus."
RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ224764; CAA12115.1; -.
DR InterPro; IPR001899; Gram_pos_anchor.
DR Pfam; PF00746; Gram_pos_anchor; 1.
DR TIGRFAMS; TIGR01167; LPXTG_anchor; 1.
DR TIGRFAMS; TIGR01168; YSIRK_signal; 1.
DR PROSITE; PS00343; GRAM_POS_ANCHORING; UNKNOWN_1.
KW SIGNAL.
FT SIGNAL 1 44 POTENTIAL.
FT CHAIN 45 913 CLUMPING FACTOR B.

SQ SEQUENCE 913 AA; 97247 MW; 38077C321F8D3E61 CRC64;

Query Match 18.0%; Score 312.5; DB 2; Length 913;
Best Local Similarity 24.9%; Pred. No. 1.9e-09;
Matches 92; Conservative 69; Mismatches 132; Indels 77; Gaps 15;

QY 6 APAAGTDITNQLTNVTVGIDSGTTVYPHQAGYVKLNLYGFSVPNSAVKGDTEKITVPKELN 65
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
208 ADAKGTNVNDKVATSNFKLEK-TTFDPNQSGNTFMAANFTVTDKVKSGDYF----- 257
66 LNGVTSTAKVPPIMAGDQVLANGVID-----SDGNVI----- 97
258 -----TAKLP-----DSLGTGNGVDYDYSNSNNTMPIADIKSTNGDVVAKATYDILTCTYT 306
98 YTFDYNVTNKDDVKATLTMPAYIDPENVKKTGNVTLATGIGSTTANKTVLVVDYK-----Y 153
: | | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
307 FVFTDYVNNKENINGQFSLPLFTDRAKAPKSGTYDANINIADENFNKKITYNYSSPIAGI 366
154 GKFNLSIKGTIDQIDKTN--NTYRQTIYVNP-----GDNVIAPVLTGNLKPNTDSNALI 207
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
367 DKPNGANISSQIIGVDTASGQNTYKQTVFVNPQKORVLGNTWV--YIKGYQDKIESSGKV 424
208 DQQNTSIKVKYKVDNAADLSESYFVNP--ENFEDVNSV--NITFPNPNQYKVEFNTPDDQ 263
425 SATDTKLRIFEVNDTSKLSDSYVADPNDSNLKEVTDQFKNRIYYEHPNVASIKFG-----D 480
264 ITTPYIVVNGHIDPNKSGDLALRSTLYGYN-----SNIIRWSMSWDNEVAFNNGSGSGDG 319
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
481 ITKTYVVLVEGHYDNTGKN--LKQVQIQENVDPVTNRDYSIFGNWNNENVVRYGGGSADG 537
320 IDKPVVPEQP 329
| | | | |
538 -DSAVNPKDP 546

RESULT 15

Q9KWX6 PRELIMINARY; PRT; 1171 AA.

ID Q9KWX6
AC Q9KWX6;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Bone sialoprotein-binding protein.
GN BBP.
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Staphylococcus.
NCBI_TaxID=1280;
[1]
SEQUENCE FROM N.A.
STRAIN=024;
RX MEDLINE=20115096; PubMed=10642520;
RA Tung H.S., Guss B., Hellman U., Persson L., Rubin K., Ryden C.;
RT "A bone sialoprotein-binding protein from Staphylococcus aureus: a member of the staphylococcal Sdr family [].";
RL Biochem. J. 345:611-619(2000).
DR EMBL; Y18653; CAB75732.1; -;
DR InterPro; IPR001899; Gram_pos_anchor.
DR Pfam; PF00746; Gram_pos_anchor; 1.
DR TIGRFAMS; TIGR01167; LPXTG_anchor; 1.
DR TIGRFAMS; TIGR01168; YSIRK_signal; 1.
DR PROSITE; PS00343; GRAM_POS_ANCHORING; UNKNOWN_1.
SQ SEQUENCE 1171 AA; 127123 MW; C5BC812F9DA5A884 CRC64;

Query Match 17.2%; Score 298; DB 2; Length 1171;
Best Local Similarity 29.2%; Pred. No. 1.6e-08;
Matches 93; Conservative 49; Mismatches 146; Indels 30; Gaps 10;

QY 2 VAADAPAAGTDITNQLT---NVTVGIDSGTTVYPHQAGYVKLNLYGFSVPNSAVKGDTEK 57
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
264 VAQPAAVASNNVNDLITVTQKMITEGIKDDGVIOAHGDEHIYTSDFKIDNAVKAGDTMT 323
58 ITVPKELNLNGVSTAKVPPIMAGD---QVLANGVIDSDGNVI-YTFDYNVTNKDDVKAT 113

Db 324 VKYDKETIPSDITD--DFTPVDITDPSGEVIAKGFEDLNTKTTTYKFTDYVDREYNVNAK 381
QY 114 LTMPAYIDPENVKKTGNVTLATGIGSTTANKTVLVVDYKYGKFNLSIKGTIDQIDKTNN 173
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
Db 382 LELNSYIDKKEVPNETNLNLTATADKETSKNVKEVQKPIVKDESNIQSIFSHLDTTKH 441
QY 174 TYRQTIYVNP---SGDNVIAPVLTGNLKP-----TDSNALIDQONTSIKVKVDNAADL 225
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
Db 442 EVEQTIYVNPPLKLNKAKNTNVTIKSGGVADNGDYTGDSIID-SNTEIKVKVVASGQQL 500
QY 226 SESYFV-NPENFEDVNSVNITFPNPNQYKVEFNTPDDQITTPYIVVNGHIDPNKSGDL 284
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
Db 501 PQSNKIYDYSQYEDVNSVTI---NKNYGTNMANINFGDIDSAYIVKVVSQYTPGAEDDL 557
QY 285 ALR-----STLYGYN 295
| | | : : : : :
Db 558 AVQQGVNMTTINKYNS 575

RESULT 16

O86488 PRELIMINARY; PRT; 1315 AA.

ID O86488
AC O86488;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Sdr protein.
GN SDR.
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Staphylococcus.
NCBI_TaxID=1280;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=NEWMAN;
RX MEDLINE=99098700; PubMed=9884231;
RA Josefsson E., McCrea K., Ni Eidhin D., O'Connell D., Cox J., Hook M.,
RA Foster T.J.;
RT "Three new members of the serine-aspartate repeat protein multigene family of Staphylococcus aureus.";
RL Microbiology 144:3387-3395(1998).
DR EMBL; AJ005646; CAA06651.1; -;
DR InterPro; IPR001899; Gram_pos_anchor.
DR Pfam; PF00746; Gram_pos_anchor; 1.
DR TIGRFAMS; TIGR01167; LPXTG_anchor; 1.
DR TIGRFAMS; TIGR01168; YSIRK_signal; 1.
SQ SEQUENCE 1315 AA; 142775 MW; 58D4E1F48EE6A689 CRC64;

Query Match 14.9%; Score 257.5; DB 2; Length 1315;
Best Local Similarity 26.9%; Pred. No. 3.2e-06;
Matches 88; Conservative 51; Mismatches 157; Indels 31; Gaps 12;

QY 13 ITNQLTNVTVGIDSGTTVYPHQAGYVKLNLYGFSVPNSAVKGDTEKITVPKELNLNGVT-- 70
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
Db 251 ITSNTTLTVVDADKNNKIVPAQ-DYLSLKSQITVDDKVKSGDYFTIKYSDTVQVYGLNPE 309
QY 71 -----STAKVPPIMAGDQVLANGVIDSDGNVI-YTFDYNVTNKDDVKATLTMPAYIDPEN 124
| | | : : : : : | | | | | : : : : : | | | | | : : : : :
Db 310 DIKNIGDIKDP---NNGETIATAKHDTANNLITYTFTDYVDRENSVQMGINYSIYMDADT 366
QY 125 VKKTGN-VTLATGIGSTTANKTVLVVDYKYGKFNLSIKG-----TIDQIDKTNT--YRQ 177
: : : | | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
Db 367 IPVSKNDVEFNVTIGNTTTKTANIQYPDYVVNNEKNSIGSAFTETVSHVGNKENPGYYKQ 426
QY 178 TIYVNPSGDNVIAPVLTGNLKPNTDSNAL--IDQNTSIKVKVDNAADLSESYFVNPN 235
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
Db 427 TIYVNPSENSLTNAKLKVQAVHSSYPNNIGQINKDVTDIKIYQPKGYTLNKGYDVNTKE 486
QY 236 FEDVNSV--NITFPNPNQYKVEFNTPDDQITTPYIVVNGHID-PNSKGDALRSTLYG 292
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
Db 487 LTDVTNQYLQKITGYGDNNNSAVIDFCNAD---SAYVVMVNTKFOYTNSSEPTLVQMATLS 542

DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Sdrc protein.
GN SDRG.
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Staphylococcus.
OX NCBI_TaxID=1280;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NEWMAN;
RX MEDLINE=99098700; PubMed=9884231;
RA Josefsen E., McCrea K., Ni Eidhin D., O'Connell D., Cox J., Hook M., Foster T.J.;
RT "Three new members of the serine-aspartate repeat protein multigene family of Staphylococcus aureus.";
RL Microbiology 144:3387-3395(1998).
DR EMBL: AJ005645; CAA06650.1; --
DR InterPro; IPR001899; Gram_pos_anchor.
DR Pfam; PF00746; Gram_pos_anchor; 1.
DR TIGRFAMs; TIGR01167; LPXTG_anchor; 1.
DR TIGRFAMs; TIGR01168; YSIRK_signal; 1.
PROSITE; PS00343; GRAM_POS_ANCHORING; UNKNOWN_1.
SQ SEQUENCE 947 AA; 102888 MW; 3C6EFD6E35121554 CRC64;

Query Match 12.2%; Score 210.5; DB 2; Length 947;
Best Local Similarity 23.1%; Pred. No. 0.00083;
Matches 83; Conservative 57; Mismatches 136; Indels 83; Gaps 16;

QY 6 APAAGTDITNQL--TNVTVGIDSG-----TTVYPHQAGYVVKLYGFSVPNSAVKGD 55
Db 178 APQOGTNVNDKVHFSNIDIAIDKGVNQTTGKTEFWATSSDVLKANYTIDDSVKEGDT 237

QY 56 FKITVPKELNLNGVTSTAKVPPIM--AGDQVLANGVIDSDGN-VIYTFDYVNTKDDVKAT 113
Db 238 FTFKYQYFRPCSVRLPSQTQNLVNAQGNIIAKGIYDSTTTNTTTFYTNVVDQYTNVRGS 297

QY 114 LTMPAYIDPEN--VKKTG---NVTLATGIGSTTANKTVLVVDYKFKYFNLSIKGTIDQI 168
Db 298 FEQVAFAKRKNATTDKTAYKMEVTL----GNDTYSEEIIVDYG-----NKKAQPLISST 347

QY 169 DKTN---TYRQTIYVNPSPGDNVIAPVLTGNL-----KPNTDSNALIDQOQNTSIKVKVD 220
Db 348 NYINNEDLSRNTAYVNPQPKNTYTKQTFVTNLTGYKFNPN-----KNFKIYEVT 397

QY 221 NAADLSESYFVNPENFEDVTSNVNITFPNPN-----QYKV-EFNTPPDQ 263
Db 398 DQNFVDSFTPDTSKLDVTDQFVYSNDNKNKTATVLDLKGQTSNKKYIIQQVAYPDNS 457

264 ITTPYIVVNGHIDPNKSGDLALRSTLYGYNSNIIWRMSWDNEVAFNNGSGSGDIDK 322
Db 458 STD-----NGKIDYTLDTD---KTKY-----SWSNSYSNVNGSSSTANGDQK 495

RESULT 20
Q9X3M7 PRELIMINARY; PRT; 1161 AA.
AC Q9X3M7;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Fibronectin-binding protein I.
GN PFBP.
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;
OC Streptococcaceae; Streptococcus.
OX NCBI_TaxID=1314;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A735;
RX MEDLINE=99270927; PubMed=10338474;
RA Rocha C.L., Fischetti V.A.;
RT "Identification and characterization of a novel fibronectin-binding protein on the surface of group A streptococci.";

RL Infect. Immun. 67:2720-2728(1999).
DR EMBL; AF071083; AAD33086.1; --
DR InterPro; IPR004237; Fn_bind.
DR InterPro; IPR001899; Gram_pos_anchor.
DR Pfam; PF02986; Fn_bind; 4.
DR PROSITE; PS00343; GRAM_POS_ANCHORING; UNKNOWN_1.
SQ SEQUENCE 1161 AA; 127425 MW; 1D1FCA1B139B5425 CRC64;

Query Match 11.8%; Score 204; DB 2; Length 1161;
Best Local Similarity 22.8%; Pred. No. 0.0024;
Matches 86; Conservative 72; Mismatches 145; Indels 74; Gaps 18;

QY 8 AAGTDITN--QLTNVTVGIDS--GTT-----VYPHQAGYVVKLYGFSVPNSAV---K 52
Db 142 AGSKDVSSSLQLENPKMSVSVSKYKTEVSSGAADFYNHAAAYFKMSPELKQDKSETINP 201

QY 53 GDTFKITVPKELNLNGVTSTAKVPPIM--AGDQVLANGVIDSDGN-VIYTFDYVNTKDD 109
Db 202 GDTFVLQDLRRLNPKGISQ--DIPKIIYDSANSPLAIGKYHAENHQLIYTFDYIAGLDK 259

QY 110 VKATLTMPAYIDPE-----NVKKT-----GNVTLATGIGSTTANKTVLVVDYE 151
Db 260 VQLSAELSLLENKEVLENTSISNFKSTIGGQEITYKGTNVNVLNGESTKESNYITNGLS 319

QY 152 KYGKFYNLSIKGTIDQIDKTNNTYRQTIYVNPSPGDNVIAPVLTGNL---KPNTDSNALI 207
Db 320 NVG-----GSIESYNTETGEFVYVYVNPRTNI--PYATMNLWGFGRRARNTSDLE 369

QY 208 DQQNTS-----IKVYKVDNAADLSESYFVNPENF---EDVTSNVNITFPNPNQYKVEFN 258
Db 370 NDANTSSAELGEIQYEVPEGEKLPSSYGVDVTKLTLRDITAGLNGFQMTKRQRIDFG 429

QY 259 TPDDQITTPYIVVNGHIDPNKSGDLALRSTLYGYNS-----NIIWRMSWDNEVAFN 311
Db 430 --NNIQNKAFIIKVTGKTQSGK-PLVVQSNLASFRGASEYAAFTPVGGNVYFQNEIALS 486

QY 312 NGSGSGDG---IDKPVV 325
Db 487 PSKSGSGSGKSEFTKPSI 503

RESULT 21
Q8RJ10 PRELIMINARY; PRT; 1161 AA.
AC Q8RJ10
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Fibronectin binding protein (Protein F2).
GN PFBP1 OR PRTF2.12.
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;
OC Streptococcaceae; Streptococcus.
OX NCBI_TaxID=1314;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A735;
RX MEDLINE=21843101; PubMed=11854196;
RA Bessen D.E., Kalia A.;
RT "Genomic localization of a T serotype locus to a recombinatorial zone encoding extracellular matrix-binding proteins in Streptococcus pyogenes.";
RL Infect. Immun. 70:1159-1167(2002).
DR EMBL; AY049088; AAL11469.1; --
DR EMBL; AF447492; AAL86412.1; --
SQ SEQUENCE 1161 AA; 127509 MW; BED4609D029B1FD7 CRC64;

Query Match 11.8%; Score 204; DB 2; Length 1161;
Best Local Similarity 22.8%; Pred. No. 0.0024;
Matches 86; Conservative 72; Mismatches 145; Indels 74; Gaps 18;

QY 8 AAGTDITN--QLTNVTVGIDS--GTT-----VYPHQAGYVVKLYGFSVPNSAV---K 52
Db 142 AGSKDVSSSLQLENPKMSVSVSKYKTEVSSGAADFYNHAAAYFKMSPELKQDKSETINP 201

Db	142	AGSKDVSSSLQLENPKMSVVSKYGKTEVSSGAADFYRNHAAAYFKMSFELKQDKSETINP	201
QY	53	GDTEFKITVPKELNLNGVTSTAKVPPIM--AGDQVLANGVIDSDGN-VIYTFTDYVNTKDD	109
Db	202	GDTEFVLQDRLRNPKGISQ--DIPKIIYDSANSPLAIGKYHAENHQLIYFTTDYIAGLDK	259
QY	110	VKATLTMPAYIDPE-----NVKKT-----GNVTLATGIGSTTANKTVLV DYE	151
Db	260	VQLSAELSLFLENKEVLENTSISNFKSTIGGQEITYKGTNVNLYGNESKESNYITNGLS	319
QY	152	KYGKFYNLSIKGTIDQIDKTNNTYRQTIYVNPBGDNVIAPVLTGNL----KPNTDSNALI	207
Db	320	NVG-----GSIESYNTETGEFVWVYVYVNPRTNI--PYATMNLWGFGRARSNTSDLE	369
QY	208	DQQNTS-----IKYVKVDNAADLSESYFVNPENF---EDVTNSVNITFPNPQYKVEFN	258
Db	370	NDANTSSAELGEIQYEVPEGEKLPSSYGVDVTKLTLRDITAGLNGCFQMTKRQRIDFG	429
QY	259	TPDDQIITPYIVVNVNGHIDPNKSGDLALRSTLYGNS-----NIIWRSMWDNEVAFN	311
QY	430	--NNIQNKAFIIKVTGKTDSGK-PLVVQSNLASFRGASEYAAFTPVGGNVYFQNEIALS	486
QY	312	NGSGSGDG---IDKPVV	325
Db	487	PSKSGSGKSEFTKPSI	503
RESULT 22			
QY	Q9KI14	PRELIMINARY; PRT; 1733 AA.	
AC	Q9KI14;		
DT	01-OCT-2000 (TrEMBLrel. 15, Created)		
DT	01-OCT-2000 (TrEMBLrel. 15, Last sequence update)		
DT	01-JUN-2002 (TrEMBLrel. 21, Last annotation update)		
DE	Putative cell-surface adhesin SdrF.		
GN	SDRF.		
OS	Staphylococcus epidermidis.		
OC	Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;		
OC	Staphylococcus.		
OX	NCBI_TaxID=1282;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=9491;		
RX	MEDLINE=20340957; PubMed=10878118;		
RA	McCrea K.W., Hartford O., Davis S., Ni Eidhin D., Lina G.,		
RA	Spezialle P., Foster T.J., Hook M.;		
RT	"The serine-aspartate repeat (Sdr) protein family in Staphylococcus epidermidis."		
QY	Microbiology 146:1535-1546(2000).		
DR	EMBL; AF245041; AAF72509.1; -		
DR	InterPro; IPR001899; Gram_pos_anchor.		
DR	Pfam; PF00746; Gram_pos_anchor; 1.		
DR	PROSITE; PS00343; GRAM_POS_ANCHORING; UNKNOWN_1.		
SQ	SEQUENCE 1733 AA; 184720 MW; D8D62EAL1692FD4E8 CRC64;		
Query Match 11.6%; Score 201.5; DB 2; Length 1733;			
Best Local Similarity 21.8%; Pred. No. 0.0055;			
Matches 79; Conservative 59; Mismatches 146; Indels 79; Gaps 15;			
QY	1	MVAADAPAAAGTDITN--QLTNVTVGID-----SGTTVPYHQAGYVKLN YGFSVPNSAVKG	53
Db	355	MSLAAEPNSGKNVNDVKITNPTLSLNKSNHANNVIWPTSNEQFNLKANYELDDSIKEG	414
QY	54	DTFKITVPKELNLNGVTSTAKVPPIMAGD-QVLANGVIDSDGN-VIYTFTDYVNTKDDVK	111
Db	415	DTFTIKYQYIRPGGLEPAIKTQLRSKGSIVANGYDKTTNTTTTFTNYVDQYQNIT	474
QY	112	ATLTMPAYIDPEN-VKKGTGNVTLATGIGSTTANKTVLV DYEKYGKFYNLSIKGTIDQIDK	170
Db	475	GSFDLIATPKRETAIKDNQNPMEVTIANEVVKKDFIVD---YGNKKONTTAAVANVDN	531
QY	171	TNNTYRQTIYVNPBGDN-----VIAPVLTGNLKPN-----TDSNALIDQQNTSIKVIK	218

Db	532	VNNKHNEVVYLNQNNQNPKYAKYFSTVKNGEFIPGEVKVYEVTDTNAMVDSFNPD L	587
QY	219	VDNAADLSESYFVNPENFEDVTNS-----VNITF-----PNPNQYKVEFNTPDDQ	263
Db	588	-----NSSNVKDVTSQFAPKVSADGTRVDINFARSMANGKKY-----	624
QY	264	ITTPYIVVNVNGHIDPNKSGDLALRSTLY---GYNSNIWRSMWDNEVAFNNGSGGDGI	320
Db	625	-----IVTQAVRPTGTGNVYTEYWLTRDGTNTNDFYRGTK-STTVTYLNGSSTAQG-	675
QY	321	DKP	323
Db	676	DNP	678
RESULT 23			
QY	P72534	PRELIMINARY; PRT; 1039 AA.	
ID	P72534		
AC	P72534;		
DT	01-FEB-1997 (TrEMBLrel. 02, Created)		
DT	01-FEB-1997 (TrEMBLrel. 02, Last sequence update)		
DT	01-DEC-2001 (TrEMBLrel. 19, Last annotation update)		
DE	Protein F2.		
GN	PRTF2.		
OS	Streptococcus pyogenes.		
OC	Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;		
OC	Streptococcaceae; Streptococcus.		
OX	NCBI_TaxID=1314;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=100076;		
RX	MEDLINE=97011581; PubMed=8858591;		
RA	Jaffe J., Natanson-Yaron S., Caparon M.G., Hanski E.;		
RT	"protein F2, a novel fibronectin-binding protein from Streptococcus pyogenes, possesses two binding domains."		
RL	Mol. Microbiol. 21:373-384(1996).		
DR	EMBL; U31980; AAC44522.1; -		
DR	InterPro; IPR004237; Fn_bind.		
DR	InterPro; IPR001899; Gram_pos_anchor.		
DR	Pfam; PF02986; Fn_bind; 4.		
DR	PROSITE; PS00343; GRAM_POS_ANCHORING; UNKNOWN_1.		
SQ	SEQUENCE 1039 AA; 115037 MW; DEC806A181A2EC3A CRC64;		
Query Match 10.5%; Score 182; DB 2; Length 1039;			
Best Local Similarity 21.9%; Pred. No. 0.035;			
Matches 81; Conservative 78; Mismatches 149; Indels 62; Gaps 18;			
QY	8	AAGTDITN--QLTNVTVGI-----DSGTTVPYHQAGYVKLN YGFSVPNSAV---K	52
Db	21	AGSKDVSSSLQLENPKMSVYKYGEQKTSNSADFYRNHAAAYFKMSFELKQDKSETINP	80
QY	53	GDTEFKITVPKELNLNGVTSTAKVPPIM--AGDQVLANGVIDSDGN-VIYTFTDYVNTKDD	109
Db	81	GDTEFVLQDRLRNPKGISQ--DIPKIIYDSSENSPLAIGKYDAKTHQLTYTFTNYIAGLDK	138
QY	110	VKATLTMPAYIDPENVKKTGNVT-LATGIG--STTANKTVLV DY---EKYGKFYN---L	159
Db	139	VQLSAELSLFLENKEVLENTNISDFKSTIGGQEITYKGTNVNLYGNESKESNYITNGLS	198
QY	160	SIKGTIDQIDKTNNTYRQTIYVNPBGDNVIAPVLTGNL-----KPNTDSNALIDQQN	211
Db	199	NVGSIESYNTETGEFVWVYVYVNPRTNIPYAVL--NLWGFAKRTAQGENDNSLVSSAQL	256
QY	212	TSIKVYKVDNAADLSESYFVNPENFEDVTNSVNITFPNP-----QYKVEFNTPDDQIT	265
Db	257	TGYDIYEVPHNYRLPTSYGVDISRL-NLRKDLEAKLPQSGTGANKRLRIDFG--ENLQG	313
QY	266	TPYIVVNVNGHIDPNKSGDLALRSTLYGYNSNIWRSMWDNEVAFNNGSGS	316
Db	314	KAFVVKVTGKADQSGK-ELIVQSHLSSFNN---WGSYKTLRPNSHVFTNEIALSPKGS	369
QY	317	GDGIDKPVVP	326

Db 370 GSGTSEFTKP 379

RESULT 24

Q8RM86

ID Q8RM86 PRELIMINARY; PRT; 1160 AA.

AC Q8RM86;

DT 01-JUN-2002 (TReMBLrel. 21, Created)

DT 01-JUN-2002 (TReMBLrel. 21, Last sequence update)

DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)

DE Fibronectin binding protein.

GN PFBPI.

OS Streptococcus pyogenes.

OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;

OC Streptococcaceae; Streptococcus.

OX NCBI_TaxID=1314;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=B737;

RX MEDLINE=21843101; PubMed=11854196;

Bessen D.E., Kalia A.;

"Genomic localization of a T serotype locus to a recombinatorial zone encoding extracellular matrix-binding proteins in Streptococcus pyogenes.";

RT Infect. Immun. 70:1159-1167(2002).

RL EMBL; AY049089; AAL11472.1; -.

DR EMBL; AY049089; AAL11472.1; -.

SQ SEQUENCE 1160 AA; 128148 MW; 097CI726A5126133 CRC64;

Query Match 10.5%; Score 182; DB 2; Length 1160;

Best Local Similarity 21.9%; Pred. No. 0.04;

Matches 81; Conservative 78; Mismatches 149; Indels 62; Gaps 18;

QY 8 AAGTDITN--QLTNVTVGI-----DSGTTVYPHQAGYVKLNYGFSVPNSAV---K 52

Db 142 AGSKDVSSSLQLENPKMSVSKYGEQKTSNSADFYRNHAAFYKMSFELKQDKSETINP 201

QY 53 GDTFKITVPKELNNGVTSTAKVPPIM--AGDQVLANGVIDSDGN-VIYFTDYVNTKDD 109

Db 202 GDTFVLQDRLRLNPKGISQ--DIPKIIYDSENSPLAIGKYDAKTHQLTYTFTNYIAGLDK 259

QY 110 VKATLTMPAYIDPENVKKTGNVT-LATGIG--STTANKTVLVDY----EKYGFYN---L 159

Db 260 VQLSAELSLFLENKEVLENTNISDFKSTIGGQEIYKGTNVNLYGNESTKESNYITNGLS 319

QY 160 SIKGTTDQIDKTNNTYRQTIYVNPBGDNVIAPIVLTGNL-----KPNTDSNALIDQON 211

Db 320 NVGGSIESYNTETGEFVYVYVNPRTNIPYAVL--NLWGFAKRTAQGENDNSLVSSAQL 377

QY 212 TSIVKVKVDNAADLSESYFVNPENFEDVTNSVNITFPNP-----QYKVEFNTPDQIT 265

Db 378 TGYDIYEVPHNYRLPTS YGVGDISRL-NLRKDLEAKLPQSTQGANRLRIDFG--ENLQG 434

QY 266 TPYIVVNGHIDPNSKGDALRSTLYGYSNIIWRS-----MSWDNEVAFNNGSGS 316

Db 435 KAFVYKVTGKADQSGK-ELIVQSHLSSFNN--WGSYKTLRPNSHVSTNEIALSPSKGS 490

QY 317 GDGIDKPVVP 326

Db 491 GSGTSEFTKP 500

RESULT 25

Q92DD5

ID Q92DD5 PRELIMINARY; PRT; 463 AA.

AC Q92DD5;

DT 01-DEC-2001 (TReMBLrel. 19, Created)

DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)

DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)

DE Hypothetical protein lin0879.

GN LIN0879.

OS Listeria innocua.

OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;

OC Listeriaceae; Listeria.

OX NCBI_TaxID=1642;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=CLIP 11262 / SEROVAR 6A;

RX PubMed=11679669;

RA Glaser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,

RA Baquero F., Berche P., Bloecker H., Brandt P., Chakraborty T.,

RA Charbit A., Chetouani F., Couve E., de Daruvar A., Dehoux P.,

RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,

RA Entian K.-D., Fsihi H., Garcia-del portillo F., Garrido P.,

RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,

RA Jones L.-M., Kaerst U., Kreft J., Kuhn M., Kunst F., Kurapkat G.,

RA Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,

RA Nordsiek G., Novella S., de Pablo B., Perez-Diaz J.-C., Purcell R.,

RA Remmel B., Rose M., Schlueter T., Simoes N., Tierrez A.,

RA Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.;

RT "Comparative genomics of Listeria species.";

RL Science 294:849-852(2001).

DR EMBL; AL596166; CAC96111.1; -.

DR Listilist; LIN00879; -.

DR InterPro; IPR001899; Gram_pos_anchor.

DR InterPro; IPR002482; LysM.

DR Pfam; PF01476; LysM; 1.

DR PROSITE; PS00343; GRAM_POS_ANCHORING; UNKNOWN_1.

KW Hypothetical protein; Complete proteome.

SQ SEQUENCE 463 AA; 49939 MW; D1C15FB01E33C07F CRC64;

Query Match 9.3%; Score 161.5; DB 16; Length 463;

Best Local Similarity 24.9%; Pred. NO. 0.17;

Matches 88; Conservative 48; Mismatches 129; Indels 89; Gaps 18;

QY 9 AGTDI-TNQLTNVTVGIDSGTTVP-HQAGYVKLNYGFSVPNSAVKGTFTKITVPKELNL 66

Db 25 AATDYGSSPFTNVALQNQNGDPVTNFKENSKVRVAYDEVITQPVVSGETMTLTIPDQLKL 84

QY 67 NGVTSTAKVPPIMAGDQVLANGVID-SDGNVIYFTDYVNTKDDVKATLTMPAYIDPENV 125

Db 85 ---INFGGFPVNDASGNTIANATIDPATGTITLTFTTHYVNTHTNLSGSLFYFNATFNSKNI 141

QY 126 KK-----TGNVTLATGIGSTTANKTVLVLDYKYGKGFYNLSIKGTI 165

Db 142 QTDQVNPFIQFPVNTTQTINTYISKVTTGGGTGSPTI-----VFKQGRM 185

QY 166 DQIDKTNNTYRQTIYVNPBGDNVIAPIVLTGNLKPNTDSNALIDQONTSIKVYKVDN 221

Db 186 D--DKDTSIIHWTVTLN----NALTPIDNAIYDTL--GTGQNLN---GNATIK-YRDAN 233

QY 222 AADLSESYFVNP-----ENFEDVTNSVNITFPNPQYKVEFNTPDQITTPYIVVNGH 275

Db 234 KKVLSSTN--VQPIALDSNRNFELTIGTLN-----NQSVV--ITYDTKITTKQKSYTN-- 281

QY 276 IDPNSKGDALRSTLYGYSNIIWRSMSWDNEVAFNNGSGSGDGIDKPVVPEQP 329

Db 282 -----KATLSGDNLDAVSRNATVND--YGGGGQGTGTPPTPPVKEEP 322

Search completed: June 23, 2003, 14:03:29

Job time : 48.8422 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 23, 2003, 13:59:31 ; Search time 8.15248 Seconds
(without alignments)
1426.837 Million cell updates/sec

Title: US-10-056-052A-20
Perfect score: 647
Sequence: 1 QVQLKESGPGLVKPSQTLTI.....YYGYDGFVYWGQGLVTVSS 121

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database : PIR_73:*
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result NO.	Score	Query Match	Length	DB ID	Description
1	534.5	82.6	117	2 S38563	Ig heavy chain V r
2	489	75.6	114	2 S11106	Ig heavy chain V r
3	489	75.6	140	2 S55028	Ig heavy chain V r
4	488	75.4	127	2 B31807	Ig heavy chain V r
5	487.5	75.3	122	2 S20809	Ig heavy chain V r
6	486.5	75.2	139	2 A32456	Ig heavy chain pre
7	482.5	74.6	122	2 A49049	Ig heavy chain V r
8	482	74.5	144	1 G2MS14	Ig heavy chain pre
9	478.5	74.0	135	2 S31913	Ig gamma-2A chain
10	477	73.7	112	2 S11100	Ig heavy chain V r
11	477	73.7	113	2 S11101	Ig heavy chain V r
12	476.5	73.6	115	2 S26470	Ig heavy chain V r
13	472	73.0	116	2 S11102	Ig heavy chain V r
14	470	72.6	140	2 S14238	Ig gamma-1 chain p
15	466.5	72.1	116	2 S42484	Ig heavy chain V r
16	465.5	71.9	141	2 S52446	Ig heavy chain V r
17	465.5	71.9	144	2 S11244	Ig heavy chain V r
18	465	71.9	121	2 D30560	Ig gamma-2a chain
19	464.5	71.8	120	2 PL0087	Ig heavy chain V r
20	464	71.7	114	2 S11099	Ig heavy chain V r
21	463	71.6	111	2 S26324	Ig heavy chain V r
22	459	70.9	115	2 S11103	Ig heavy chain V r
23	457.5	70.7	118	2 S32786	Ig heavy chain (an
24	457	70.6	115	2 S11107	Ig heavy chain V r
25	452	69.9	110	2 S26323	Ig heavy chain V r
26	452	69.9	118	2 PQ0266	Ig heavy chain V r
27	451	69.7	112	2 S11108	Ig heavy chain V r
28	450	69.6	231	2 PC4155	Ig gamma-2b chain
29	448.5	69.3	107	2 S14492	Ig heavy chain V r

30	447.5	69.2	107	2 S14493	Ig heavy chain V r
31	447	69.1	106	2 S26322	Ig heavy chain V r
32	443	68.5	109	2 PH1026	Ig heavy chain V r
33	443	68.5	112	2 S11098	Ig heavy chain V r
34	441.5	68.2	107	2 S14491	Ig heavy chain V r
35	441.5	68.2	110	2 PH1024	Ig heavy chain V r
36	440	68.0	121	2 S33131	Ig heavy chain V r
37	439.5	67.9	106	2 S14489	Ig heavy chain V r
38	437	67.5	117	2 S10111	Ig heavy chain V r
39	435	67.2	116	2 PL0198	anti-DNA autoantib
40	434	67.1	115	1 HVMS14	Ig heavy chain pre
41	433	66.9	109	2 PH1025	Ig heavy chain V r
42	433	66.9	114	2 S26321	Ig heavy chain V r
43	433	66.9	114	2 S11104	Ig heavy chain V r
44	431.5	66.7	107	2 S14506	Ig heavy chain V r
45	428	66.2	114	2 S11105	Ig heavy chain V r
46	426	65.8	109	2 S11109	Ig heavy chain V r
47	426	65.8	116	2 A33932	Ig mu chain precu
48	424	65.5	130	2 S31690	Ig heavy chain V r
49	424	65.5	140	2 I37782	Ig variable region
50	423	65.4	100	2 S14490	Ig heavy chain V r
51	423	65.4	116	1 GLMS10	Ig heavy chain pre
52	421	65.1	100	2 A25913	Ig heavy chain pre
53	419	64.8	140	2 S54245	Ig mu heavy chain
54	418	64.6	135	2 S78051	Ig heavy chain pre
55	417	64.5	101	2 S03466	Ig heavy chain V r
56	416.5	64.4	129	2 S44114	Ig heavy chain V r
57	414.5	64.1	130	2 S30534	Ig heavy chain V r
58	414	64.0	155	2 S31511	Ig heavy chain - h
59	413.5	63.9	118	2 S24443	Ig heavy chain V r
60	413.5	63.9	137	2 S31676	Ig heavy chain V r
61	410	63.4	155	2 S31512	Ig heavy chain - h
62	407.5	63.0	95	2 S17605	Ig heavy chain V r
63	405	62.6	147	2 S13519	Ig heavy chain V r
64	400.5	61.9	139	2 S31586	Ig heavy chain V r
65	400.5	61.9	470	2 S22080	Ig heavy chain pre
66	399.5	61.7	141	2 S54226	Ig mu heavy chain
67	398	61.5	109	2 PH1023	Ig heavy chain V r
68	398	61.5	144	2 S54244	Ig mu heavy chain
69	396.5	61.3	135	2 B31933	Ig heavy chain pre
70	396	61.2	121	2 S44113	Ig heavy chain V r
71	394.5	61.0	145	2 S78055	Ig heavy chain pre
72	394	60.9	90	2 A49042	Ig heavy chain V r
73	394	60.9	132	2 S54235	Ig mu heavy chain
74	393.5	60.8	134	2 B24672	Ig heavy chain pre
75	393.5	60.8	141	2 S54236	Ig mu heavy chain
76	392	60.6	149	2 S30752	Ig heavy chain pre
77	388.5	60.0	140	2 S78052	Ig heavy chain pre
78	388	60.0	97	2 S55372	Ig heavy chain V r
79	388	60.0	139	2 S31696	Ig heavy chain V r
80	387	59.8	98	2 S21812	Ig heavy chain V r
81	387	59.8	138	2 S54246	Ig mu heavy chain
82	387	59.8	140	2 A49045	Ig heavy chain V r
83	387	59.8	140	2 S54239	Ig mu heavy chain
84	387	59.8	142	2 S54241	Ig mu heavy chain
85	386.5	59.7	122	2 S69909	Ig V-D-J region (M
86	386	59.7	112	2 PH1022	Ig heavy chain V r
87	385.5	59.6	117	2 B33989	Ig heavy chain V-2
88	385	59.5	123	2 S30530	Ig heavy chain V r
89	385	59.5	134	2 S54906	Ig heavy chain V r
90	384.5	59.4	116	2 S38718	Ig heavy chain V r
91	384.5	59.4	137	2 S54234	Ig mu heavy chain
92	384.5	59.4	143	2 S54249	Ig mu heavy chain
93	384	59.4	140	2 S54227	Ig mu heavy chain
94	383.5	59.3	111	2 PH1019	Ig heavy chain V r
95	381	58.9	119	2 C53285	Ig heavy chain V a
96	381	58.9	121	2 S37200	Ig heavy chain V r
97	379	58.6	146	2 S09711	Ig heavy chain V r
98	378.5	58.5	136	2 S07637	Ig heavy chain V r
99	377.5	58.3	118	2 S20780	Ig heavy chain V r
100	377.5	58.3	122	2 S69912	Ig V-D-J region (N

QY 69 ISKDNSKNQVFLKMNSLTAADTAVYYCARKGEFYIG-YDGFVYWGQGLTVTVSS 121
IIIIII:IIIIII III IIII III I: IIIII IIIII
Db 61 ISKDNSKSQVFLKMNSLQDDTARYICAR---VYGYNEAMDYGQGSTVTVSS 111

RESULT 22
S11103
Ig heavy chain V region (clone NQ5-61.1.2) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 19-Mar-1997 #sequence_revision 23-Aug-1997 #text_change 21-Jan-2000
C:Accession: S11103
R;Kaartinen, M.; Griffiths, G.M.; Markham, A.F.; Milstein, C.
Nature 304, 320-324, 1983
A:Title: mRNA sequences define an unusually restricted IgG response to 2-phenyloxazolone
A:Reference number: S07331; MUID:83271467; PMID:6877353
A:Accession: S11103
A:Molecule type: mRNA
A:Residues: 1-115 <NAT>
C:Superfamily: immunoglobulin V region; immunoglobulin homology
F:15-97/Domain: immunoglobulin homology <IMM>

Query Match 70.9%; Score 459; DB 2; Length 115;
Best Local Similarity 73.6%; Pred. No. 9.5e-36;
Matches 89; Conservative 9; Mismatches 15; Indels 8; Gaps 1;

QY 1 QVQLKESGPGLVKPSQTLTITCTISGFSLSRYSVHWVRQPPGKGLWLGMIWGGNTDYN 60
IIIIIIIIII III:IIIIII:II II: IIIIIIIIIIIIIIIII:II I:I:II
Db 1 QVQLKESGPGLVAPSQSLTITCTVSGXSLTSXGVHWVRQPPGKGLWLGVIWAGGSTNYN 60

QY 61 SALKSRLSISKDNSKNQVFLKMNSLTAADTAVYYCARKGEFYIGYDGFVYWGQGLTVTVS 120
III IIIIIIIIIII: IIIIIII III:IIIIII I III IIIIIII
Db 61 SALMSRLSISKDNSKSVFLKMNSLQDDTAMYYICAR-----DRGAYWGXGTLTVTVS 112

QY 121 S 121
Db 113 A 113

RESULT 23
S32786
Ig heavy chain (anti-biotin) - mouse
C:Species: Mus musculus (house mouse)
C:Date: 02-Dec-1993 #sequence_revision 10-Nov-1995 #text_change 23-Jul-1999
C:Accession: S32786
R;Bagci, H.; Kohen, F.; Kuscuoğlu, U.; Bayer, E.A.; Wilchek, M.
Proc. Natl. Acad. Sci. USA 90, 12222-12226, 1993
A:Title: Monoclonal anti-biotin antibodies simulate avidin in the recognition of biotin.
A:Reference number: S32786; MUID:93245957; PMID:8482366
A:Accession: S32786
A:Status: preliminary
A:Molecule type: nucleic acid
A:Residues: 1-118 <BAG>
A:Cross-references: GB:S59639; NID:g2999965; PIDN:AAB26438.1; PID:g2999966
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:15-97/Domain: immunoglobulin homology <IMM>

Query Match 70.7%; Score 457.5; DB 2; Length 118;
Best Local Similarity 72.1%; Pred. No. 1.3e-35;
Matches 88; Conservative 11; Mismatches 18; Indels 5; Gaps 2;

QY 1 QVQLKESGPGLVKPSQTLTITCTISGFSLSRYSVHWVRQPPGKGLWLGMIWGGNTDYN 60
II:IIIIII III:IIIIII:IIII: I IIIIIIIIIIIIIIIII:IIII I:II
Db 1 QVKLQESGPGLVAPSQSLTITCTVSGFSLTAYGVDWVRQPPGKGLWLGVIWGGRTNYN 60

QY 61 SALKSRLSISKDNSKNQVFLKMNSLTAADTAVYYCARKGEFYIGYD-GFVYWGQGLTVTV 119
I I IIIII IIIII:IIII III III: : I I IIIII III
Db 61 SGLMSRLSIRKDNSKSVFLTMNSLQDDTAKYYCVK---HTNWDGCFAYWGQGTTVTV 116

QY 120 SS 121
II
Db 117 SS 118

RESULT 24
S11107
Ig heavy chain V region (clone NQ2-45.10.4) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 19-Mar-1997 #sequence_revision 18-Jul-1997 #text_change 21-Jan-2000
C:Accession: S11107
R;Kaartinen, M.; Griffiths, G.M.; Markham, A.F.; Milstein, C.
Nature 304, 320-324, 1983
A:Title: mRNA sequences define an unusually restricted IgG response to 2-phenyloxazolone
A:Reference number: S07331; MUID:83271467; PMID:6877353
A:Accession: S11107
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-115 <KAA>
C:Superfamily: immunoglobulin V region; immunoglobulin homology
F:15-97/Domain: immunoglobulin homology <IMM>

Query Match 70.6%; Score 457; DB 2; Length 115;
Best Local Similarity 71.9%; Pred. No. 1.5e-35;
Matches 87; Conservative 11; Mismatches 15; Indels 8; Gaps 1;

QY 1 QVQLKESGPGLVKPSQTLTITCTISGFSLSRYSVHWVRQPPGKGLWLGMIWGGNTDYN 60
IIIIII: III I: II: IIIII:IIII: I IIIIIIIIIIIIIIIII:II I:I:II
Db 1 QVQLKQXGPGXVQXSQXSITCTVSGFSLTSGYVHWVRQPPGKGLWLGVIWGGSTDYN 60

QY 61 SALKSRLSISKDNSKNQVFLKMNSLTAADTAVYYCARKGEFYIGYDGFVYWGQGLTVTVS 120
: I IIIIIIIIIII:III IIII I III:IIII :I IIIII IIIII
Db 61 AAFISRLSISKDNSKSQVFFKMNSLQANDTAIYYICAR-----NGGDYWGQGSTVTVS 112

QY 121 S 121
Db 113 S 113

RESULT 25
S26323
Ig heavy chain V region - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 19-Mar-1998 #sequence_revision 19-Mar-1998 #text_change 20-Jun-2000
C:Accession: S26323
R;Stark, S.E.; Caton, A.J.
J. Exp. Med. 174, 613-624, 1991
A:Title: Antibodies that are specific for a single amino acid interchange in a protein
A:Reference number: S26309; MUID:91341421; PMID:1908510
A:Accession: S26323
A:Molecule type: mRNA
A:Residues: 1-110 <STA>
A:Cross-references: EMBL:X59184; NID:g52061; PIDN:CAA41894.1; PID:g1334033
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:6-87/Domain: immunoglobulin homology <IMM>

Query Match 69.9%; Score 452; DB 2; Length 110;
Best Local Similarity 78.6%; Pred. No. 4e-35;
Matches 88; Conservative 6; Mismatches 16; Indels 2; Gaps 1;

QY 10 GLVKPSQTLTITCTISGFSLSRYSVHWVRQPPGKGLWLGMIWGGNTDYN SALKSRLSI 69
III III:IIIIII:IIII: I I:IIIIIIIIIIIIIIII:IIII I:IIIIIIIIIIII
Db 1 GLVAPSQSLTITCTVSGFSLTGYGVNVRQPPGKGLWLGMIWGGSTDYN SALKSRLSI 60

QY 70 SKDNSKNQVFLKMNSLTAADTAVYYCARKGEFYIGYDGFVYWGQGLTVTVSS 121
IIIIII:IIIIIIII III IIIII I III IIIII IIIII
Db 61 SKDNSKSQVFLKMNSLQDDTARYICARVP--YGNVDAMDYWGQGSTVTVSS 110

Search completed: June 23, 2003, 14:04:15
Job time : 9.15248 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 23, 2003, 13:59:02 ; Search time 4.50532 Seconds
(without alignments)
1113.936 Million cell updates/sec

Title: US-10-056-052A-20
Perfect score: 647
Sequence: 1 QVQLKESGPGLVKPSQTL SI.....YGYDGFVYWGQGLTVTVSS 121

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

ched: 112892 seqs, 41476328 residues 112892
1 number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	482	74.5	144	1 HV43_MOUSE	P01819 mus musculus
2	434	67.1	115	1 HV44_MOUSE	P01820 mus musculus
3	423	65.4	116	1 HV45_MOUSE	P01821 mus musculus
4	396.5	61.3	135	1 HV02_XENLA	P20957 xenopus lae
5	374.5	57.9	122	1 HV3G_HUMAN	P01768 homo sapien
6	370.5	57.3	146	1 HV2I_HUMAN	P06331 homo sapien
7	365	56.4	137	1 HV46_MOUSE	P01822 mus musculus
8	349.5	54.0	122	1 HV3H_HUMAN	P01769 homo sapien
9	349	53.9	121	1 HV3J_HUMAN	P01771 homo sapien
10	346	53.5	117	1 HV2G_HUMAN	P01825 homo sapien
11	344	53.2	136	1 HV01_XENLA	P20956 xenopus lae
12	342	52.9	119	1 HV2C_HUMAN	P01816 homo sapien
13	338.5	52.3	116	1 HV61_MOUSE	P01832 mus musculus
14	338	52.2	129	1 HV2F_HUMAN	P01824 homo sapien
15	335.5	51.9	126	1 HV3K_HUMAN	P01772 homo sapien
16	333	51.5	113	1 HV47_MOUSE	P01823 mus musculus
17	333	51.5	119	1 HV3I_HUMAN	P01770 homo sapien
18	332.5	51.4	119	1 HV40_MOUSE	P01810 mus musculus
19	329.5	50.9	136	1 HV16_MOUSE	P01783 mus musculus
20	328.5	50.8	119	1 HV37_MOUSE	P01807 mus musculus
21	325.5	50.3	114	1 HV3B_HUMAN	P01763 homo sapien
22	325.5	50.3	126	1 HV2A_HUMAN	P01814 homo sapien
23	324.5	50.2	119	1 HV38_MOUSE	P01808 mus musculus
24	322	49.8	118	1 HV39_MOUSE	P01809 mus musculus
25	322	49.8	142	1 HV01_RAT	P01805 rattus norv
26	318.5	49.2	116	1 HV60_MOUSE	P18531 mus musculus
27	317.5	49.1	120	1 HV2B_HUMAN	P01815 homo sapien
28	316.5	48.9	116	1 HV3T_HUMAN	P01781 homo sapien
29	316	48.8	115	1 HV3D_HUMAN	P01765 homo sapien
30	316	48.8	123	1 HV25_MOUSE	P01794 mus musculus
31	314.5	48.6	122	1 HV3A_HUMAN	P01762 homo sapien
32	314	48.5	117	1 HV17_MOUSE	P01786 mus musculus
33	312	48.2	125	1 HV2D_HUMAN	P01817 homo sapien

34	311	48.1	116	1	HV05_CARAU	P19181 carassius a
35	310.5	48.0	120	1	HV3U_HUMAN	P01782 homo sapien
36	310	47.9	119	1	HV3L_HUMAN	P01773 homo sapien
37	309	47.8	115	1	HV32_MOUSE	P01801 mus musculus
38	308.5	47.7	114	1	HV2A_RABIT	P01827 oryctolagus
39	308	47.6	117	1	HV62_MOUSE	P18533 mus musculus
40	308	47.6	120	1	HV3E_HUMAN	P01766 homo sapien
41	307.5	47.5	139	1	HV07_MOUSE	P01751 mus musculus
42	305	47.1	117	1	HV2B_RABIT	P01828 oryctolagus
43	304	47.0	113	1	HV30_MOUSE	P01799 mus musculus
44	303.5	46.9	122	1	HV20_MOUSE	P01789 mus musculus
45	303.5	46.9	136	1	HV2C_RABIT	P01829 oryctolagus
46	303.5	46.9	147	1	HV2H_HUMAN	P04438 homo sapien
47	303	46.8	115	1	HV3F_HUMAN	P01767 homo sapien
48	302.5	46.8	117	1	HV3C_HUMAN	P01764. homo sapien
49	301	46.5	117	1	HV02_CANFA	P01785 canis famil
50	301	46.5	123	1	HV18_MOUSE	P01787 mus musculus
51	301	46.5	123	1	HV19_MOUSE	P01788 mus musculus
52	300	46.4	115	1	HV33_MOUSE	P01802 mus musculus
53	300	46.4	138	1	HV48_MOUSE	P03980 mus musculus
54	299	46.2	113	1	HV27_MOUSE	P01796 mus musculus
55	299	46.2	117	1	HV41_MOUSE	P01811 mus musculus
56	299	46.2	123	1	HV22_MOUSE	P01791 mus musculus
57	298.5	46.1	118	1	HV51_MOUSE	P06330 mus musculus
58	298	46.1	113	1	HV29_MOUSE	P01798 mus musculus
59	296	45.7	113	1	HV31_MOUSE	P01800 mus musculus
60	295	45.6	121	1	HV2E_HUMAN	P01818 homo sapien
61	295	45.6	123	1	HV23_MOUSE	P01792 mus musculus
62	293.5	45.4	119	1	HV3M_HUMAN	P01774 homo sapien
63	293.5	45.4	122	1	HV21_MOUSE	P01790 mus musculus
64	293	45.3	113	1	HV28_MOUSE	P01797 mus musculus
65	292	45.1	123	1	HV24_MOUSE	P01793 mus musculus
66	290.5	44.9	119	1	HV3N_HUMAN	P01775 homo sapien
67	288.5	44.6	114	1	HV01_CANFA	P01784 canis famil
68	288	44.5	117	1	HV42_MOUSE	P01812 mus musculus
69	288	44.5	140	1	HV02_MOUSE	P01746 mus musculus
70	287.5	44.4	119	1	HV3P_HUMAN	P01777 homo sapien
71	287	44.4	121	1	HV01_MOUSE	P01745 mus musculus
72	287	44.4	144	1	HV26_MOUSE	P01795 mus musculus
73	285.5	44.1	111	1	HV35_MOUSE	P01804 mus musculus
74	285.5	44.1	116	1	HV1A_RABIT	P01826 oryctolagus
75	285	44.0	120	1	HV03_MOUSE	P01747 mus musculus
76	283	43.7	117	1	HV12_MOUSE	P01756 mus musculus
77	281.5	43.5	120	1	HV50_MOUSE	P06329 mus musculus
78	281	43.4	113	1	HV34_MOUSE	P01803 mus musculus
79	280	43.3	117	1	HV13_MOUSE	P01757 mus musculus
80	279.5	43.2	117	1	HV3O_HUMAN	P01776 homo sapien
81	279.5	43.2	137	1	HV11_MOUSE	P01755 mus musculus
82	278	43.0	136	1	HV15_MOUSE	P01759 mus musculus
83	276.5	42.7	117	1	HV55_MOUSE	P18526 mus musculus
84	276.5	42.7	118	1	HV3V_HUMAN	P80419 homo sapien
85	274	42.3	97	1	HV56_MOUSE	P18527 mus musculus
86	271.5	42.0	98	1	HV57_MOUSE	P18528 mus musculus
87	270.5	41.8	117	1	HV54_MOUSE	P18525 mus musculus
88	269	41.6	116	1	HV3R_HUMAN	P01779 homo sapien
89	268.5	41.5	124	1	HV1D_HUMAN	P01760 homo sapien
90	266.5	41.2	116	1	HV36_MOUSE	P01806 mus musculus
91	265.5	41.0	117	1	HV14_MOUSE	P01758 mus musculus
92	264.5	40.9	117	1	HV59_MOUSE	P18530 mus musculus
93	264	40.8	125	1	HV1F_HUMAN	P06326 homo sapien
94	262	40.5	116	1	HV3Q_HUMAN	P01778 homo sapien
95	261.5	40.4	117	1	HV52_MOUSE	P06327 mus musculus
96	261.5	40.4	117	1	HV58_MOUSE	P18529 mus musculus
97	256.5	39.6	117	1	HV04_MOUSE	P01748 mus musculus
98	256.5	39.6	147	1	HVIC_HUMAN	P01744 homo sapien
99	255.5	39.5	117	1	HV03_CARAU	P19180 carassius a
100	252.5	39.0	117	1	HV1B_HUMAN	P01743 homo sapien

ALIGNMENTS

	Matches	69;	Conservative	22;.	Mismatches	25;	Indels	6;	Gaps	2;
QY		1	QVQLKESGPGLVKPSQTLSITCTISGFSLRSYSVHWVRQP	PGKGLEWLGMIWG	GGNTDYN	60				
			: : : :	: :	:	:	:	:	:	:
Dd		1	QVLEQSGPGLVRPSQTLSLTCTVSGSTFSDYYTWVRQP	PGRGLEWIGYVFYHGT	SDDT	60				
			: : : :	: :	:	:	:	:	:	:
QY		61	SALKSRLSISKDNSSKNQVFLKMNSLTAADTAVYVCARKGEFY	YGVDGV-YWGQG	TIVTV	119				
			: :: :::	:	:	:				
Dd		61	TPLRSRVTMLVDTSKNQFSRLSSVTAAATAVIYCARN----	LIAGCIDVWGQS	LTVTV	115				
QY		120	SS	121						
Dd		116	SS	117						

```

RESULT 11
HV01_XENLA
ID HV01_XENLA STANDARD; PRT; 136 AA.
AC P20956;
DT 01-FEB-1991 (Rel. 17; Created)
01-FEB-1991 (Rel. 17; Last sequence update)
15-JUL-1999 (Rel. 38; Last annotation update)
IG heavy chain V region XIG8 precursor (Fragment).
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88176921; PubMed=2451244;
RA Schwager J., Mikoryak C.A., Steiner L.A.;
RT "Amino acid sequence of heavy chain from Xenopus laevis IgM deduced
RT from cDNA sequence: implications for evolution of immunoglobulin
RT domains.";
RL Proc. Nati. Acad. Sci. U.S.A. 85:2245-2249(1988).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/
CC or send an email to license@isb-sib.ch).
CC -----

```

RESULT 12.

```

HV2C_HUMAN
ID HV2C_HUMAN STANDARD; PRT; 119 AA.
AC P01816;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig heavy chain V-II region DAW.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE.
RX MEDLINE=70258837; PubMed=5449120;
RA Press E.M., Hogg N.M.;
RT "The amino acid sequences of the Fd fragments of two human gamma-1
RT heavy chains."
RL Biochem. J. 117:641-660(1970).
CC -!- MISCELLANEOUS; THIS CHAIN WAS OBTAINED FROM IGG1 ISOLATED FROM THE
CC SERUM OF A PATIENT WITH HYPERGAMMAGLOBULINEMIA.
DR PIR; A02091; G1HUDW.
DR HSSP; P01789; IMCP.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; ig; 1.
DR SMART; SM00406; IGV; 1.
KW Immunoglobulin v region.
FT MOD_RES 1
FT NON_TER 119
SQ SEQUENCE 119 AA; 13045 MW; 4E13E00214BAD789 CRC64;

Query Match 52.9%; Score 342; DB 1; Length 119;
Best Local Similarity 57.6%; Pred. NO. 1.9e-28;
Matches 72; Conservative 15; Mismatches 28; Indels 10; Gaps 4;

QY 1 QVQLKESGPGLVKPSQTLISITCTISGFSLS--RYSVHVVRQPPGKGLEWLGMIWGGGNTD 58
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 QVTLRESGPAIVRPQTTLTCTFSGFSLSGETMCMVAIRQPPGGALEWLA--WDILNDD 58

QY 59 --YNSALKRSLISIKDKNQVFLKMNSLTAADTAVYVCARKGEFYGYDGFVYWGQGT 116
   |::|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 59 KYYGASLETRLAVSKDTSKNQVLSMNTVGPQDTATYYCARS---CGSQYFDYWGQGIL 114

QY 117 VTVSS 121
   |||||
Db 115 VTVSS 119

RESULT 13
HV61_MOUSE
ID HV61_MOUSE STANDARD; PRT; 116 AA.
AC P18532;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig heavy chain V region IB43 precursor.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=BALB/CJ;
RC MEDLINE=89279149; PubMed=2499654;
RA Levy N.S., Malipiero U.V., Lebecque S.G., Gearhart P.J.;
RT "Early onset of somatic mutation in immunoglobulin VH genes during
RT the primary immune response."
RL J. Exp. Med. 169:2007-2019(1989).
CC -!- MISCELLANEOUS; THIS SEQUENCE BELONGS TO THE VH3660 SUBFAMILY.
DR PIR; JT0508; HVMS1B.
DR HSSP; P01825; 7FAB.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.

```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 23, 2003, 13:59:10 ; Search time 15.6613 Seconds
(without alignments)
1591.926 Million cell updates/sec

Title: US-10-056-052A-20
Perfect score: 647
Sequence: 1 QVQLKESGPGLVKPSQTL SI.....YYGYDGFVYWGQGLVTVSS 121

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

671580 seqs, 206047115 residues
number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

- Database : SPTREMBL_21:*
- 1: sp_archaea:*
 - 2: sp_bacteria:*
 - 3: sp_fungi:*
 - 4: sp_human:*
 - 5: sp_invertebrate:*
 - 6: sp_mammal:*
 - 7: sp_mhc:*
 - 8: sp_organelle:*
 - 9: sp_phage:*
 - 10: sp_plant:*
 - 11: sp_rodent:*
 - 12: sp_virus:*
 - 13: sp_vertibrate:*
 - 14: sp_unclassified:*
 - 15: sp_rvirus:*
 - 16: sp_bacteriap:*
 - 17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	465.5	71.9	121	11 Q99NG4	Q99ng4 mus musculu
2	463.5	71.6	482	11 Q91X92	Q91x92 mus musculu
3	413	63.8	119	4 Q9UL73	Q9ul73 homo sapien
4	402.5	62.2	613	4 Q96EY0	Q96ey0 homo sapien
5	385.5	59.6	496	4 Q96KX8	Q96kx8 homo sapien
6	373.5	57.7	150	4 Q95973	Q95973 homo sapien
7	369	57.0	479	11 Q99M22	Q99m22 mus musculu
8	363.5	56.2	573	4 Q8WU38	Q8wu38 homo sapien
9	363	56.1	121	4 Q9UL96	Q9ul96 homo sapien
10	359	55.5	121	4 Q9UL71	Q9ul71 homo sapien
11	355.5	54.9	613	4 Q8WUK1	Q8wuk1 homo sapien
12	354	54.7	588	4 Q8WUX4	Q8wux4 homo sapien
13	354	54.7	597	4 Q9BU10	Q9bu10 homo sapien
14	354	54.7	618	4 Q96AA6	Q96aa6 homo sapien
15	351	54.3	116	4 Q9UL93	Q9ul93 homo sapien
16	350.5	54.2	122	4 Q9UL75	Q9ul75 homo sapien

17	350.5	54.2	147	4 Q9Y509	Q9y509 homo sapien
18	350	54.1	597	4 Q9BQB8	Q9bqb8 homo sapien
19	350	54.1	597	4 Q96BB9	Q96bb9 homo sapien
20	346.5	53.6	469	11 Q8R3V9	Q8r3v9 mus musculu
21	345.5	53.4	473	4 Q8TC63	Q8tc63 homo sapien
22	344	53.2	113	4 Q9UL90	Q9ul90 homo sapien
23	343.5	53.1	112	4 Q9HCC1	Q9hcc1 homo sapien
24	337.5	52.2	118	4 Q9UL72	Q9ul72 homo sapien
25	337	52.1	119	11 Q920E7	Q920e7 mus musculu
26	336	51.9	473	11 Q9D8L4	Q9d8l4 mus musculu
27	333	51.5	131	4 Q9UL88	Q9ul88 homo sapien
28	332.5	51.4	122	4 Q9UL84	Q9ul84 homo sapien
29	331.5	51.2	487	11 Q99KA4	Q99ka4 mus musculu
30	330.5	51.1	494	4 Q96K68	Q96k68 homo sapien
31	328.5	50.8	471	4 Q8TC77	Q8tc77 homo sapien
32	324.5	50.2	298	11 Q9QYF0	Q9qyf0 mus musculu
33	319.5	49.4	118	4 Q9UL91	Q9ul91 homo sapien
34	318.5	49.2	484	11 Q8VEA0	Q8vea0 mus musculu
35	317.5	49.1	124	6 Q9N0W6	Q9n0w6 oryctolagus
36	316.5	48.9	124	6 Q9N0W4	Q9n0w4 oryctolagus
37	316	48.8	486	11 Q91207	Q91207 mus musculu
38	315.5	48.8	463	11 Q99LC4	Q99lc4 mus musculu
39	314.5	48.6	473	11 Q99L25	Q99l25 mus musculu
40	312.5	48.3	473	11 Q91205	Q91205 mus musculu
41	311	48.1	481	11 Q91WT1	Q91wt1 mus musculu
42	308	47.6	97	4 Q43234	Q43234 homo sapien
43	307.5	47.5	145	11 Q924Q7	Q924q7 mus musculu
44	307	47.4	142	11 Q924Q1	Q924q1 mus musculu
45	306.5	47.4	124	4 Q9UL92	Q9ul92 homo sapien
46	306.5	47.4	488	11 Q91WR1	Q91wr1 mus musculu
47	306	47.3	104	4 Q9UL87	Q9ul87 homo sapien
48	305.5	47.2	145	11 Q924P7	Q924p7 mus musculu
49	303	46.8	147	11 Q925S3	Q925s3 mus musculu
50	303	46.8	479	11 Q91WP5	Q91wp5 mus musculu
51	301.5	46.6	145	11 Q924R1	Q924r1 mus musculu
52	299.5	46.3	145	11 Q924Q6	Q924q6 mus musculu
53	299	46.2	146	11 Q924R8	Q924r8 mus musculu
54	298.5	46.1	143	11 Q924Q5	Q924q5 mus musculu
55	298.5	46.1	145	11 Q924R4	Q924r4 mus musculu
56	298.5	46.1	613	11 Q8VCX7	Q8vcx7 mus musculu
57	298	46.1	123	11 Q8VIJ1	Q8vij1 mus musculu
58	298	46.1	489	11 Q8VCX4	Q8vcx4 mus musculu
59	297.5	46.0	481	11 Q8VCV5	Q8vcv5 mus musculu
60	297	45.9	117	11 Q9QXF0	Q9qxf0 mus musculu
61	296	45.7	146	11 Q924Q3	Q924q3 mus musculu
62	295.5	45.7	143	11 Q924R0	Q924r0 mus musculu
63	295.5	45.7	145	11 Q924Q9	Q924q9 mus musculu
64	295.5	45.7	159	4 Q96QS0	Q96qs0 homo sapien
65	295	45.6	117	11 Q9QXE9	Q9qxe9 mus musculu
66	294.5	45.5	143	11 Q924Q0	Q924q0 mus musculu
67	294	45.4	120	11 Q920E8	Q920e8 mus musculu
68	293.5	45.4	139	11 Q924R5	Q924r5 mus musculu
69	293.5	45.4	437	11 Q9R1A4	Q9ria4 mus musculu
70	292	45.1	144	11 Q924P5	Q924p5 mus musculu
71	291	45.0	140	11 Q924R2	Q924r2 mus musculu
72	290	44.8	118	4 Q9UL74	Q9ul74 homo sapien
73	290	44.8	140	11 Q924P8	Q924p8 mus musculu
74	290	44.8	168	11 Q8VDC9	Q8vdc9 mus musculu
75	290	44.8	474	11 Q8R3H6	Q8r3h6 mus musculu
76	290	44.8	484	11 Q99LA6	Q99la6 mus musculu
77	288.5	44.6	143	11 Q924R7	Q924r7 mus musculu
78	288.5	44.6	143	11 Q91VA2	Q91va2 mus musculu
79	288.5	44.6	278	11 Q921K1	Q921k1 mus musculu
80	287	44.4	468	11 Q99L31	Q99l31 mus musculu
81	287	44.4	480	11 Q91XE1	Q91xe1 mus musculu
82	287	44.4	497	4 Q8WY24	Q8wy24 homo sapien
83	286.5	44.3	137	11 Q924R6	Q924r6 mus musculu
84	286	44.2	119	5 Q9GY22	Q9gy22 schistosoma
85	285.5	44.1	141	11 Q924Q4	Q924q4 mus musculu
86	285.5	44.1	143	11 Q924P6	Q924p6 mus musculu
87	285.5	44.1	145	11 Q924R3	Q924r3 mus musculu
88	285	44.0	117	11 Q921C6	Q921c6 mus musculu
89	285	44.0	142	11 Q924Q2	Q924q2 mus musculu

Db 1 QVQLQESGPGLVKPSQTLTCTVSGGSICSYVWSWIRQPPGKGLEWIGYIYSGSTNYT 60

QY 61 SALKSRLSISKDNQVFLKMNLSLTAADTAVYYCARKGEF--YYGYDGFVYWGQGLTVT 118

Db 61 PSLKSRVTISVDRSKNQFSLKLTSLTAADTAVYFCARLSNWGPYY----FDYWGQGLTVT 116

QY 119 VSS 121

Db 117 VSS 119

RESULT 4

Q96EY0

ID Q96EY0 PRELIMINARY; PRT; 613 AA.

AC Q96EY0;

DT 01-DEC-2001 (TrEMBLrel. 19, Created)

DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)

DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)

Unknown (protein for MGC:20337).

Homo sapiens (Human).

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=B-CELL;

RA Strausberg R.;

RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; BC011857; AAH11857.1; -

DR InterPro; IPR003598; Ig_c2.

DR InterPro; IPR003006; Ig_MHC.

DR Pfam; PF00047; Ig; 5.

DR SMART; SM00408; IGC2; 2.

DR PROSITE; PS00290; IG_MHC; UNKNOWN_3.

KW Immunoglobulin domain.

SQ SEQUENCE 613 AA; 67273 MW; 31214203FB8421E7 CRC64;

Query Match 62.2%; Score 402.5; DB 4; Length 613;

Best Local Similarity 64.5%; Pred. No. 3.8e-36;

Matches 78; Conservative 19; Mismatches 23; Indels 1; Gaps 1;

QY 1 QVQLKESGPGLVKPSQTLTCTISGFSLSRYSVHWVRRQPPGKGLEWLGMIWGGNTDYN 60

Db 20 QVQLQESGPGLVKPSQTLTCTVSGGSISYYWSWIRQPPAGKGLEWIGRIYTSNSTYN 79

QY 61 SALKSRLSISKDNQVFLKMNLSLTAADTAVYYCARKGEFYGYDGFVYWGQGLTVTVS 120

80 PSLKSRVTMSVDTSKNQFSLKLSLSSVTAADTAVYYCASQ-PWELPTVGLFYWGQGLTVTVS 138

121 S 121

139 S 139

Db

RESULT 5

Q96KX8

ID Q96KX8 PRELIMINARY; PRT; 496 AA.

AC Q96KX8;

DT 01-DEC-2001 (TrEMBLrel. 19, Created)

DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)

DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)

DE Hypothetical 53.4 kDa protein.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=LUNG;

RA Strausberg R.;

RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; BC016369; AAH16369.1; -

DR InterPro; IPR003006; Ig_MHC.

DR Pfam; PF00047; ig; 4.

DR PROSITE; PS00290; IG_MHC; UNKNOWN_1.

KW Hypothetical protein.

SQ SEQUENCE 496 AA; 53391 MW; D346929849040D69 CRC64;

Query Match 59.6%; Score 385.5; DB 4; Length 496;

Best Local Similarity 62.9%; Pred. No. 2.3e-34;

Matches 78; Conservative 18; Mismatches 25; Indels 3; Gaps 2;

QY 1 QVQLKESGPGLVKPSQTLTCTISGFSLSRYSVH--WVRQPPGKGLEWLGMIWGGNTD 58

Db 20 QLQLQESGPGLVKSSQTLTCTVSGGSISSSSYVWGWIROPKGLEWIANITYSGITY 79

QY 59 YNSALKSRLSISKDNQVFLKMNLSLTAADTAVYYCARKGEFYGYDGFV-YWGQGLTV 117

Db 80 YNPSLKSRVTISVDTSKNQFSLKVRSVTAADTAVYFCARHGYSRSGRTGAIDYWGQGLTV 139

QY 118 TVSS 121

Db 140 TVSS 143

RESULT 6

O95973

ID O95973 PRELIMINARY; PRT; 150 AA.

AC O95973;

DT 01-MAY-1999 (TrEMBLrel. 10, Created)

DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)

DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)

DE VH4 heavy chain variable region precursor (Fragment).

GN IGM.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RA Suh C.-H., Song C.-H., Lee C.-H., Lee S.-K.;

RT "Clonal proliferation of IgM secreting B cell in the synovium of Behcet's patient with arthritis.";

RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF103795; AAC79084.1; -

DR HSSP; P01825; 7FAB.

DR InterPro; IPR003006; Ig_MHC.

DR InterPro; IPR003596; Ig_v.

DR Pfam; PF00047; ig; 1.

DR SMART; SM00406; IGV; 1.

KW Signal.

FT SIGNAL 1 19 POTENTIAL.

FT CHAIN 20 >150 VH4 HEAVY CHAIN VARIABLE REGION.

FT NON_TER 150 150

SQ SEQUENCE 150 AA; 16315 MW; 85664E04938AA7C9 CRC64;

Query Match 57.7%; Score 373.5; DB 4; Length 150;

Best Local Similarity 59.3%; Pred. No. 1.1e-33;

Matches 73; Conservative 20; Mismatches 23; Indels 7; Gaps 2;

QY 1 QVQLKESGPGLVKPSQTLTCTISGFSLS--RYSVHWVRRQPPGKGLEWLGMIWGGNTD 58

Db 20 QLQLQESGPGLVKPSQTLTCTVSGGSISSTNYVWGWIROPKGLEWIGSLHNSGSDY 79

QY 59 YNSALKSRLSISKDNQVFLKMNLSLTAADTAVYYCARKGEFYGYDGFVYWGQGLTVT 118

Db 80 YNPSLKSRVTISVDTSKNQFSLRLSSVTAADTAVYYCAR-----LGMGAFDFWGHGTMVT 134

QY 119 VSS 121

Db 135 VSS 137

RESULT 7

Q99M22

ID Q99M22 PRELIMINARY; PRT; 479 AA.

AC Q99M22;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Hypothetical 52.0 kDa protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC02091; AAH02091.1; -
DR HSSP; P01810; 2FBJ.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR003597; Ig_cl.
DR InterPro; IPR003600; Ig_like.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 4.
DR SMART; SM00409; IG; 3.
DR SMART; SM00407; IGcl; 3.
DR SMART; SM00406; IGV; 1.
DR SMART; SM00410; IG_like; 1.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_2.
KW Hypothetical protein.
SQ SEQUENCE 479 AA; 51992 MW; 768E39A138918892 CRC64;

Query Match 57.0%; Score 369; DB 11; Length 479;
Best Local Similarity 58.7%; Pred. No. 1.5e-32;
Matches 71; Conservative 22; Mismatches 22; Indels 6; Gaps 2;

QY 2 VOLKESGPGLVKPSQTLSTITCTISGFSL-SRYSVHVVRQPPGKGLEWLGMIWGSGNTDYN 60
Db 20 VQLQESGPGLVKPSQSLTCSVTGYSTSGYYWNWIRQPPGNKLEWMGYINYDCSNYN 79
QY 61 SALKSRLSISKDNSKNOVFLKMNLSLTAAADTAVYYCARKGEFYGYDGFVYWGQGLVTVS 120
Db 80 PSLKNRISITRDTSKNQFFLKLNSVTETDATTYYCASR----GYSWFPNWGQGLVTVS 134
QY 121 S 121
Db 135 A 135

RESULT 8
Q8WU38 PRELIMINARY; PRT; 573 AA.
Q8WU38;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Hypothetical 63.0 kDa protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=TONSIL;
RA Strausberg R.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC021276; AAH21276.1; -
DR InterPro; IPR003599; Ig.
DR InterPro; IPR003597; Ig_cl.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 4.
DR SMART; SM00409; IG; 1.
DR SMART; SM00407; IGcl; 3.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_2.

KW Hypothetical protein.
SQ SEQUENCE 573 AA; 62967 MW; FD072344033AC530 CRC64;

Query Match 56.2%; Score 363.5; DB 4; Length 573;
Best Local Similarity 53.5%; Pred. No. 7.6e-32;
Matches 68; Conservative 23; Mismatches 27; Indels 9; Gaps 3;

QY 1 QVQLKESGPGLVKPSQTLSTITCTISGFSLSRYSVHVVRQPPGKGLEWLGMIWGSGNTDY 59
Db 20 EVQLVESGGGLVQPGRLSLRSLCAASGFTDDYAMHWVRQAPGKGLEWVSGISWNSGSIGY 79
QY 60 NSALKSRLSISKDNSKNOVFLKMNLSLTAAADTAVYYCARKGE-----FYYGYDGFVYWGQ 114
Db 80 ADSVKGRFTISRDNKNSLYLQMNLSRAEDTALYYCAKHGSGSYIGYIYGM---VWGQ 136
QY 115 TLVTVSS 121
Db 137 TTVTVSS 143

RESULT 9
Q9UL96 PRELIMINARY; PRT; 121 AA.
AC Q9UL96;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Myosin-reactive immunoglobulin heavy chain variable region
DE (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98277139; PubMed=9614934;
RA Wu X., Liu B., Van der Merwe P.L., Kalis N.N., Berney S.M.,
RA Young D.C.;
RT "Myosin-reactive autoantibodies in rheumatic carditis and normal fetus.";
RL Clin. Immunol. Immunopathol. 87:184-192(1998).
DR EMBL; AF035018; AAD56254.1; -
DR HSSP; P01825; 7FAB.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IGV; 1.
DR NON_TER 1
FT NON_TER 121 121
SQ SEQUENCE 121 AA; 13695 MW; D582D450596BDD35 CRC64;

Query Match 56.1%; Score 363; DB 4; Length 121;
Best Local Similarity 60.0%; Pred. No. 1.2e-32;
Matches 75; Conservative 15; Mismatches 27; Indels 8; Gaps 3;

QY 1 QVQLKESGPGLVKPSQTLSTITCTISGFSL--SRYSVHVVRQPPGKGLEWLGMIWGSGNTD 58
Db 1 QITLKESGPTLVKPTQTTLTCTFSGFSLTTSGLMDVGVIRQPPGKALEWLALIYWDGDKR 60
QY 59 YNSALKSRLSISKDNSKNOVFLKMNLSLTAAADTAVYYCA--RKGEFYGYDGFVYWGQGL 116
Db 61 YSPSLKSRLTITKDTSKNQVDLTMTFMDPWTATTYYCAHRKSGDGY----FDYWGQGL 116
QY 117 VTVSS 121
Db 117 VTVSS 121

RESULT 10
Q9UL71 PRELIMINARY; PRT; 121 AA.
AC Q9UL71;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)

DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DE 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Myosin-reactive immunoglobulin heavy chain variable region
DE (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98277139; PubMed=9614934;
RA Wu X., Liu B., Van der Merwe P.L., Kallis N.N., Berney S.M.,
RA Young D.C.;
RT "Myosin-reactive autoantibodies in rheumatic carditis and normal
RT fetus.";
RL Clin. Immunol. Immunopathol. 87:184-192(1998).
DR EMBL; AF035043; AAD56279.1; -.
DR HSSP; P01772; 2FB4.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IGV; 1.
DR NON_TER 1
FT NON_TER 121
FT NON_TER 121
SQ SEQUENCE 121 AA; 13154 MW; 2F045CCFA5D50736 CRC64;

Query Match 55.5%; Score 359; DB 4; Length 121;
Best Local Similarity 57.4%; Pred. No. 3.5e-32;
Matches 70; Conservative 21; Mismatches 29; Indels 2; Gaps 2;

QY 1 QVQLKESGPGLVKPSQTLSTITCTISGFSLSRYSVHWVRQPPGKGLWLGMIWG-GGNTDY 59
Db 1 EVQLVESGGGVQPGGSLRLFCASGFTFDGYAMHWVRQAPGKGLWVSLISGDDGGSTYY 60

QY 60 NSALKSRLSISKDNSKNQVFLKMNSLTAADTAVYYCARKGEFYGYDGFVYWGQGLTVTV 119
Db 61 ADSVKGRFTISRDNKNSLYLQMNLSRAEDTALYYCA-KGKVTTIYDRFDIWGQGTMTV 119

QY 120 SS 121
Db 120 SS 121

RESULT 11
Q8WUK1 PRELIMINARY; PRT; 613 AA.
ID Q8WUK1
Q8WUK1;
01-MAR-2002 (TrEMBLrel. 20, Created)
01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Hypothetical 67.3 kDa protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=TONSIL;
RA Strausberg R.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DDBJ databases.
DR EMBL; BC020240; AAH20240.1; -.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR003597; Ig_cl.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 5.
DR SMART; SM00409; IG; 2.
DR SMART; SM00407; IGcl; 4.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_3.
KW Hypothetical protein.
SQ SEQUENCE 613 AA; 67296 MW; 60C7F5950671E315 CRC64;

Query Match 54.9%; Score 355.5; DB 4; Length 613;
Best Local Similarity 55.7%; Pred. No. 6.4e-31;
Matches 68; Conservative 23; Mismatches 28; Indels 3; Gaps 2;

QY 1 QVQLKESGPGLVKPSQTLSTITCTISGFSLSRYSVHWVRQPPGKGLWLGMI-WGGGNTDY 59
Db 20 QVQLVESGGGVQPGGSLRLSCAASGFTSSYGMHWVRQAPGKGLWVAISYDGSNKYY 79

QY 60 NSALKSRLSISKDNSKNQVFLKMNSLTAADTAVYYCARKGEFYGYDGFVYWGQGLTVTV 119
Db 80 ADSVKGRFTISRDNKNSLYLQMNLSRAEDTAVYYCAK--DWSEGVETFDIWGQGTMTV 137

QY 120 SS 121
Db 138 SS 139

RESULT 12
Q8WUX4 PRELIMINARY; PRT; 588 AA.
ID Q8WUX4
Q8WUX4;
01-MAR-2002 (TrEMBLrel. 20, Created)
01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Hypothetical 64.4 kDa protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LYMPH;
RA Strausberg R.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DDBJ databases.
DR EMBL; BC019235; AAH19235.1; -.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR003597; Ig_cl.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 5.
DR SMART; SM00409; IG; 2.
DR SMART; SM00407; IGcl; 4.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_3.
KW Hypothetical protein.
SQ SEQUENCE 588 AA; 64438 MW; FC60DBAD82B39FD7 CRC64;

Query Match 54.7%; Score 354; DB 4; Length 588;
Best Local Similarity 57.0%; Pred. No. 8.9e-31;
Matches 73; Conservative 18; Mismatches 27; Indels 10; Gaps 2;

QY 1 QVQLKESGPGLVKPSQTLSTITCTISGFSLSRYSVHWVRQPPGKGLWLGMIWGNTDYN 60
Db 20 QVQLQGWGAGLLKPSSETLSLTCGVYGGSFSGYYSWIRQPPGKGLWIGEINHSSTNYN 79

QY 61 SALKSRLSISKDNSKNQVFLKMNSLTAADTAVYYCAR-----KGEFYGYDGFVYWGQ 113
Db 80 PSLKSRVTISVDTSKKQLSLKLSVNAADTAVYYCARVITRASPGTDGRYGM---VWGQ 136

QY 114 GTLVTVSS 121
Db 137 GTTVTVSS 144

RESULT 13
Q9BU10 PRELIMINARY; PRT; 597 AA.
ID Q9BU10
Q9BU10;
01-JUN-2001 (TrEMBLrel. 17, Created)
01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Hypothetical 65.3 kDa protein.
OS Homo sapiens (Human).

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 23, 2003, 14:00:06 ; Search time 7.07979 Seconds
(without alignments)
502.864 Million cell updates/sec

Title: US-10-056-052A-20
Perfect score: 647
Sequence: 1 QVQLKESGPGLVKPSQFLSI.....YYGYDGFVWGQGLTVTVSS 121

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues
al number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. NO. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	506	78.2	119	1	US-07-634-278-60	Sequence 60, Appl
2	506	78.2	119	1	US-08-477-728-60	Sequence 60, Appl
3	506	78.2	119	1	US-08-474-040-60	Sequence 60, Appl
4	506	78.2	119	1	US-08-487-200-60	Sequence 60, Appl
5	506	78.2	119	4	US-08-484-537-60	Sequence 60, Appl
6	506	78.2	138	1	US-07-634-278-33	Sequence 33, Appl
7	506	78.2	138	1	US-08-477-728-33	Sequence 33, Appl
8	506	78.2	138	1	US-08-474-040-33	Sequence 33, Appl
9	506	78.2	138	1	US-08-487-200-33	Sequence 33, Appl
10	506	78.2	138	4	US-08-484-537-33	Sequence 33, Appl
11	502	77.6	113	2	US-08-606-293-6	Sequence 6, Appli
12	494.5	76.4	137	2	US-08-621-751A-8	Sequence 8, Appli
13	492	76.0	119	2	US-08-752-844-16	Sequence 16, Appl
14	492	76.0	119	2	US-08-591-196-16	Sequence 16, Appl
15	491.5	76.0	242	6	5455030-15	Patent No. 5455030
16	487.5	75.3	120	2	US-08-652-558-8	Sequence 8, Appli
17	487.5	75.3	120	4	US-09-254-189-5	Sequence 5, Appli
18	487.5	75.3	239	2	US-08-860-174A-2	Sequence 2, Appli
19	487	75.3	113	2	US-08-606-293-2	Sequence 2, Appli
20	486.5	75.2	120	2	US-08-652-558-38	Sequence 38, Appl
21	486	75.1	222	2	US-08-190-199A-67	Sequence 67, Appl
22	486	75.1	235	2	US-08-190-199A-61	Sequence 61, Appl
23	481.5	74.4	137	2	US-08-621-751A-4	Sequence 4, Appli
24	479.5	74.1	120	2	US-08-652-558-7	Sequence 7, Appli
25	479.5	74.1	120	4	US-09-254-189-4	Sequence 4, Appli
26	471.5	72.9	116	3	US-08-397-411-3	Sequence 3, Appli
27	471.5	72.9	121	3	US-08-881-037-67	Sequence 67, Appl

28	471.5	72.9	273	3	US-08-397-411-6	Sequence 6, Appli
29	471.5	72.9	446	3	US-08-397-411-7	Sequence 7, Appli
30	468	72.3	116	2	US-08-308-494A-21	Sequence 21, Appl
31	467	72.2	119	1	US-08-667-769A-15	Sequence 15, Appl
32	467	72.2	119	1	US-08-667-769A-62	Sequence 62, Appl
33	467	72.2	119	5	PCT-US95-17082A-15	Sequence 15, Appl
34	467	72.2	119	5	PCT-US95-17082A-62	Sequence 62, Appl
35	466.5	72.1	120	2	US-08-652-558-4	Sequence 4, Appli
36	466.5	72.1	120	2	US-08-652-558-5	Sequence 5, Appli
37	466.5	72.1	120	4	US-09-254-189-2	Sequence 2, Appli
38	466.5	72.1	140	4	US-08-943-136-4	Sequence 4, Appli
39	466.5	72.1	140	4	US-08-973-518-4	Sequence 4, Appli
40	465.5	71.9	239	6	5455030-13	Patent No. 5455030
41	464.5	71.8	120	2	US-08-652-558-6	Sequence 6, Appli
42	464.5	71.8	120	4	US-09-254-189-3	Sequence 3, Appli
43	463	71.6	119	1	US-08-467-420A-62	Sequence 62, Appl
44	463	71.6	119	1	US-08-470-110A-62	Sequence 62, Appl
45	463	71.6	119	2	US-08-940-371-62	Sequence 62, Appl
46	459.5	71.0	120	2	US-08-652-558-9	Sequence 9, Appli
47	459.5	71.0	120	4	US-09-254-189-6	Sequence 6, Appli
48	458.5	70.9	120	4	US-08-057-430A-26	Sequence 26, Appl
49	458	70.8	152	2	US-08-752-844-4	Sequence 4, Appli
50	458	70.8	152	2	US-08-591-196-4	Sequence 4, Appli
51	458	70.8	152	4	US-09-192-838B-4	Sequence 4, Appli
52	458	70.8	263	2	US-08-752-844-66	Sequence 66, Appl
53	455.5	70.4	107	1	US-07-942-245-14	Sequence 14, Appl
54	455.5	70.4	120	2	US-08-353-372A-36	Sequence 36, Appl
55	455	70.3	119	3	US-08-483-749A-14	Sequence 14, Appl
56	454.5	70.2	112	4	US-09-189-129-3	Sequence 3, Appli
57	452.5	69.9	116	3	US-08-397-411-4	Sequence 4, Appli
58	449	69.4	264	4	US-08-564-164A-4	Sequence 4, Appli
59	447.5	69.2	111	2	US-08-470-139-6	Sequence 6, Appli
60	447.5	69.2	111	4	US-09-347-061-6	Sequence 6, Appli
61	447.5	69.2	111	4	US-09-537-911A-67	Sequence 67, Appl
62	447	69.1	119	1	US-07-634-278-37	Sequence 37, Appl
63	447	69.1	119	1	US-07-634-278-61	Sequence 61, Appl
64	447	69.1	119	1	US-08-477-728-37	Sequence 37, Appl
65	447	69.1	119	1	US-08-477-728-61	Sequence 61, Appl
66	447	69.1	119	1	US-08-474-040-37	Sequence 37, Appl
67	447	69.1	119	1	US-08-474-040-61	Sequence 61, Appl
68	447	69.1	119	1	US-08-487-200-37	Sequence 37, Appl
69	447	69.1	119	1	US-08-487-200-61	Sequence 61, Appl
70	447	69.1	119	4	US-08-484-537-37	Sequence 37, Appl
71	447	69.1	119	4	US-08-484-537-61	Sequence 61, Appl
72	437	67.5	97	3	US-08-881-037-66	Sequence 66, Appl
73	434	67.1	115	3	US-08-881-037-22	Sequence 22, Appl
74	431	66.6	119	4	US-09-025-769B-39	Sequence 39, Appl
75	431	66.6	119	4	US-09-025-769B-65	Sequence 65, Appl
76	429	66.3	111	1	US-08-467-420A-15	Sequence 15, Appl
77	429	66.3	111	1	US-08-470-110A-15	Sequence 15, Appl
78	429	66.3	111	2	US-08-940-371-15	Sequence 15, Appl
79	429	66.3	111	3	US-08-637-647-15	Sequence 15, Appl
80	427.5	66.1	215	4	US-09-170-769A-6	Sequence 6, Appli
81	426	65.8	119	1	US-08-467-420A-19	Sequence 19, Appl
82	426	65.8	119	1	US-08-470-110A-19	Sequence 19, Appl
83	426	65.8	119	1	US-08-667-769A-19	Sequence 19, Appl
84	426	65.8	119	2	US-08-940-371-19	Sequence 19, Appl
85	426	65.8	119	3	US-08-637-647-19	Sequence 19, Appl
86	426	65.8	119	5	PCT-US95-17082A-19	Sequence 19, Appl
87	423	65.4	107	1	US-08-122-546-12	Sequence 12, Appl
88	423	65.4	107	2	US-08-764-938-12	Sequence 12, Appl
89	423	65.4	107	3	US-09-131-052-12	Sequence 12, Appl
90	423	65.4	107	4	US-09-131-053A-12	Sequence 12, Appl
91	421.5	65.1	144	2	US-08-116-778E-36	Sequence 36, Appl
92	421.5	65.1	144	2	US-08-438-562-36	Sequence 36, Appl
93	421.5	65.1	144	2	US-08-483-528B-100	Sequence 100, App
94	420.5	65.0	118	4	US-09-025-769B-25	Sequence 25, Appl
95	419.5	64.8	135	2	US-08-470-139-28	Sequence 28, Appl
96	419.5	64.8	135	4	US-09-347-061-28	Sequence 28, Appl
97	411.5	63.6	122	1	US-08-360-125-11	Sequence 11, Appl
98	411.5	63.6	122	2	US-08-450-578-11	Sequence 11, Appl
99	411.5	63.6	122	2	US-09-017-628-11	Sequence 11, Appl
100	411.5	63.6	122	2	US-09-014-880-11	Sequence 11, Appl

ALIGNMENTS

RESULT 1
US-07-634-278-60
; Sequence 60, Application US/07634278
; Patent No. 5530101
; GENERAL INFORMATION:
; APPLICANT: QUEEN, Cary L.
; APPLICANT: CO, Man Sung
; APPLICANT: SCHNEIDER, William P.
; APPLICANT: LANDOLFI, Nicholas F.
; APPLICANT: COELINGH, Kathleen L.
; APPLICANT: SELICK, Harold E.
; TITLE OF INVENTION: IMPROVED HUMANIZED IMMUNOGLOBLINS
; NUMBER OF SEQUENCES: 113
; CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend Khourie and Crew
STREET: 379 Lytton Avenue
CITY: Palo Alto
STATE: California
COUNTRY: US
ZIP: 94301

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/634,278
FILING DATE: 19-DEC-1990
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/590,274
FILING DATE: 28-SEP-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/310,252
FILING DATE: 13-FEB-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/290,975
FILING DATE: 28-DEC-1988
ATTORNEY/AGENT INFORMATION:
NAME: Smith, William M
REGISTRATION NUMBER: 30,223
REFERENCE/DOCKET NUMBER: 11823-002600
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 326-2400
TELEFAX: (415) 326-2422
INFORMATION FOR SEQ ID NO: 60:
SEQUENCE CHARACTERISTICS:
LENGTH: 119 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-07-634-278-60

Query Match 78.2%; Score 506; DB 1; Length 119;
Best Local Similarity 75.2%; Pred. No. 4.5e-44;
Matches 91; Conservative 16; Mismatches 12; Indels 2; Gaps 1;
QY 1 QVQLKESGPGLVKPSQTLSTICTISGFSLSRYSVHWVRQPPGKGLWLGMIWGGGNTDYN 60
Db 1 QVQLKQSGPGLVQPSSQSLSTICTVSGFSVTSYGVHWIRQSPGKGLWLGVIWGGGSTDYN 60
QY 61 SALKSRSLTSKDNKSNQVFLKMNLSLTAAADTAVYYCARKGEFYGYDGFVYWGQGLTVTVS 120
Db 61 AAFISRLTISKDNKSKQVFFKVNLSLPADTAIYYCARAGD--YNYDGFAYWGQGLTVTVS 118
QY 121 S 121

Db 119 A 119
RESULT 2
US-08-477-728-60
; Sequence 60, Application US/08477728
; Patent No. 5585089
; GENERAL INFORMATION:
; APPLICANT: QUEEN, Cary L.
; APPLICANT: SCHNEIDER, William P.
; APPLICANT: SELICK, Harold E.
; TITLE OF INVENTION: IMPROVED HUMANIZED IMMUNOGLOBLINS
; NUMBER OF SEQUENCES: 113
; CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: Palo Alto
STATE: California
COUNTRY: US
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/477,728
FILING DATE: 07-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/634,278
FILING DATE: 19-DEC-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/590,274
FILING DATE: 28-SEP-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/310,252
FILING DATE: 13-FEB-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/290,975
FILING DATE: 28-DEC-1988
ATTORNEY/AGENT INFORMATION:
NAME: Smith, William M
REGISTRATION NUMBER: 30,223
REFERENCE/DOCKET NUMBER: 11823-002600
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 326-2400
TELEFAX: (415) 326-2422
INFORMATION FOR SEQ ID NO: 60:
SEQUENCE CHARACTERISTICS:
LENGTH: 119 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-477-728-60

Query Match 78.2%; Score 506; DB 1; Length 119;
Best Local Similarity 75.2%; Pred. No. 4.5e-44;
Matches 91; Conservative 16; Mismatches 12; Indels 2; Gaps 1;
QY 1 QVQLKESGPGLVKPSQTLSTICTISGFSLSRYSVHWVRQPPGKGLWLGMIWGGGNTDYN 60
Db 1 QVQLKQSGPGLVQPSSQSLSTICTVSGFSVTSYGVHWIRQSPGKGLWLGVIWGGGSTDYN 60
QY 61 SALKSRSLTSKDNKSNQVFLKMNLSLTAAADTAVYYCARKGEFYGYDGFVYWGQGLTVTVS 120
Db 61 AAFISRLTISKDNKSKQVFFKVNLSLPADTAIYYCARAGD--YNYDGFAYWGQGLTVTVS 118
QY 121 S 121
Db 119 A 119

US-08-477-728-33
: Sequence 33, Application US/08477728
: Patent No. 5585089
: GENERAL INFORMATION:
: APPLICANT: QUEEN, Cary L.
: APPLICANT: SCHNEIDER, William P.
: APPLICANT: SELICK, Harold E.
: TITLE OF INVENTION: IMPROVED HUMANIZED IMMUNOGLOBULINS
: NUMBER OF SEQUENCES: 113
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Townsend and Townsend and Crew LLP
: STREET: Two Embarcadero Center, 8th Floor
: CITY: Palo Alto
: STATE: California
: COUNTRY: US
: ZIP: 94111
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/477,728
: FILING DATE: 07-JUN-1995
: CLASSIFICATION: 424
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/634,278
: FILING DATE: 19-DEC-1990
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/590,274
: FILING DATE: 28-SEP-1990
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/310,252
: FILING DATE: 13-FEB-1989
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/290,975
: FILING DATE: 28-DEC-1988
: ATTORNEY/AGENT INFORMATION:
: NAME: Smith, William M
: REGISTRATION NUMBER: 30,223
: REFERENCE/DOCKET NUMBER: 11823-002600
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (415) 326-2400
: TELEFAX: (415) 326-2422
: INFORMATION FOR SEQ ID NO: 33:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 138 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
US-08-477-728-33

```

Query Match      78.2%; Score 506; DB 1; Length 138;
Best Local Similarity 75.2%; Pred. No. 5.4e-44;
Matches 91; Conservative 16; Mismatches 12; Indels 2; Gaps 1;

QY      1 QVQLKESGPGLVKPSQTL SITCTISGFSLSRYSVHVRQPPGKGLEWLGMIWGGGNTDYN 60
         |||||:|||||:|||||:|||||: |||||: |||||: |||||: |||||: |||||:
Db      20 QVQLKQSGPGLVQPQSLSITCTVSGFSVTSYGVHWIRQSPGKGLEWLGVIWGGGSTDYN 79

QY      61 SALKSRLLSISKDKNQVFLKMNLSLTAADTAVYYCARKGEFFYYGYDGFVYWGQGLTIVTS 120
         : |||||:|||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||:
Db      80 AAFISRLTISKDKNKSQVFFKVNLSLPADTAIIYCARAGD--YNYDGFAYWGQGLTIVTS 137

QY      121 S 121
         :
Db      138 A 138

```

RESULT 8
US-08-474-040-33
; Sequence 33, Application US/08474040
; Patent No. 5693761

GENERAL INFORMATION:
APPLICANT: QUEEN, Cary L.
APPLICANT: CO, Man Sung
APPLICANT: SCHNEIDER, William P.
APPLICANT: LANDOLFI, Nicholas F.
APPLICANT: COELINGH, Kathleen L.
APPLICANT: SELICK, Harold E.
TITLE OF INVENTION: IMPROVED HUMANIZED IMMUNOGLOBULINS
NUMBER OF SEQUENCES: 113
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend Khourie and Crew
STREET: 379 Lytton Avenue
CITY: Palo Alto
STATE: California
COUNTRY: US
ZIP: 94301
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/474,040
FILING DATE: 07-JUN-1995
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/634,278
FILING DATE: 19-DEC-1990
APPLICATION NUMBER: US 07/590,274
FILING DATE: 28-SEP-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/310,252
FILING DATE: 13-FEB-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/290,975
FILING DATE: 28-DEC-1988
ATTORNEY/AGENT INFORMATION:
NAME: Smith, William M
REGISTRATION NUMBER: 30,223
REFERENCE/DOCKET NUMBER: 11823-002600
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 326-2400
TELEFAX: (415) 326-2422
INFORMATION FOR SEQ ID NO: 33:
SEQUENCE CHARACTERISTICS:
LENGTH: 138 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein

US-08-474-040-33

```

Query Match      78.2%; Score 506; DB 1; Length 138;
Best Local Similarity 75.2%; Pred. NO. 5.4e-44;
Matches 91; Conservative 16; Mismatches 12; Indels 2; Gaps 1;

QY      1 QVQLKESGPGLVKPSQTL SITCTISGFSLSRYSVHVVRRQPPGKGLEWLGMIWGGSNTDYN 60
      |||||:|||||:|||||:|||||:| ||||:| |||||:| |||||:| |||||:|
Db      20 QVQLKQSGPGLVQPSQSL SITCTVSGFSVTSYGVHVIIRQSPGKGLEWLGVIWGGSSTDYN 79
      |||||:|||||:|||||:|||||:| ||||:| |||||:| |||||:| |||||:|

QY      61 SALKSRLSISKDNSKNQVFLKMNSLTAADTAVYYCARKGEFFYYGYDGFVYWGQGLTVTS 120
      : ||||:|||||:|||||:||||| |||||:|||||:| |||||:|||||:| |||||:|
Db      80 AAFISRLTISKDNSKSQVFFKVNLSLPADTAIIYICARAGD--YNYDGFAYWGQGLTVTS 137
      |||||:|||||:|||||:||||| |||||:|||||:| |||||:|||||:| |||||:|

QY      121 S 121
      :
Db      138 A 138

```

RESULT 9
US-08-487-200-33
; Sequence 33, Application US/08487200
; Patent No. 5693762
; GENERAL INFORMATION:

APPLICANT: QUEEN, Cary L.
APPLICANT: CO, Man Sung
APPLICANT: SCHNEIDER, William P.
APPLICANT: LANDOLFI, Nicholas F.
APPLICANT: COELINGH, Kathleen L.
APPLICANT: SELICK, Harold E.
TITLE OF INVENTION: IMPROVED HUMANIZED IMMUNOGLOBULINS
NUMBER OF SEQUENCES: 113
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew
STREET: 379 Lytton Avenue
CITY: Palo Alto
STATE: California
COUNTRY: US
ZIP: 94301
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/487,200
FILING DATE: 7-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/634,278
FILING DATE: 19-DEC-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/590,274
FILING DATE: 28-SEP-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/310,252
FILING DATE: 13-FEB-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/290,975
FILING DATE: 28-DEC-1988
ATTORNEY/AGENT INFORMATION:
NAME: Smith, William M
REGISTRATION NUMBER: 30,223
REFERENCE/DOCKET NUMBER: 11823-002610
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 326-2400
TELEFAX: (415) 326-2422
INFORMATION FOR SEQ ID NO: 33:
SEQUENCE CHARACTERISTICS:
LENGTH: 138 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-487-200-33

Query Match 78.2%; Score 506; DB 1; Length 138;
Best Local Similarity 75.2%; Pred. No. 5.4e-44;
Matches 91; Conservative 16; Mismatches 12; Indels 2; Gaps 1;
QY 1 QVQLKESGPGLVKPSQTLSTICTISGFSLSRYSVHWVRQPPGKGLEWLGMIWGSGGNTDYN 60
Db 20 QVQLKQSGPGLVQPSQSLSTICTVSGFSVTSYGVHWIRQSPGKGLEWLGVIWSGGSTDYN 79
QY 61 SALKSRLSISKDNKNQVFLKMNLSLTAADTAVIYCARKGEFYGYDGFVYWGQGTLTVTS 120
Db 80 AAFISRLTISKDNKSKQVFFKVNLSQPADTAIYYCARAGD--YNYDGFAYWGQGTLTVTS 137
QY 121 S 121
Db 138 A 138

RESULT 10
US-08-484-537-33
; Sequence 33, Application US/08484537
; Patent No. 6180370
; GENERAL INFORMATION:

APPLICANT: QUEEN, Cary L.
APPLICANT: CO, Man Sung
APPLICANT: SCHNEIDER, William P.
APPLICANT: LANDOLFI, Nicholas F.
APPLICANT: COELINGH, Kathleen L.
APPLICANT: SELICK, Harold E.
TITLE OF INVENTION: IMPROVED HUMANIZED IMMUNOGLOBULINS
NUMBER OF SEQUENCES: 113
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend Khourie and Crew
STREET: 379 Lytton Avenue
CITY: Palo Alto
STATE: California
COUNTRY: US
ZIP: 94301
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,537
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/634,278
FILING DATE: 19-DEC-1990
APPLICATION NUMBER: US 07/590,274
FILING DATE: 28-SEP-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/310,252
FILING DATE: 13-FEB-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/290,975
FILING DATE: 28-DEC-1988
ATTORNEY/AGENT INFORMATION:
NAME: Smith, William M
REGISTRATION NUMBER: 30,223
REFERENCE/DOCKET NUMBER: 11823-002600
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 326-2400
TELEFAX: (415) 326-2422
INFORMATION FOR SEQ ID NO: 33:
SEQUENCE CHARACTERISTICS:
LENGTH: 138 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-484-537-33

Query Match 78.2%; Score 506; DB 4; Length 138;
Best Local Similarity 75.2%; Pred. No. 5.4e-44;
Matches 91; Conservative 16; Mismatches 12; Indels 2; Gaps 1;
QY 1 QVQLKESGPGLVKPSQTLSTICTISGFSLSRYSVHWVRQPPGKGLEWLGMIWGSGGNTDYN 60
Db 20 QVQLKQSGPGLVQPSQSLSTICTVSGFSVTSYGVHWIRQSPGKGLEWLGVIWSGGSTDYN 79
QY 61 SALKSRLSISKDNKNQVFLKMNLSLTAADTAVIYCARKGEFYGYDGFVYWGQGTLTVTS 120
Db 80 AAFISRLTISKDNKSKQVFFKVNLSQPADTAIYYCARAGD--YNYDGFAYWGQGTLTVTS 137
QY 121 S 121
Db 138 A 138

RESULT 11
US-08-606-293-6
; Sequence 6, Application US/08606293
; Patent No. 5874082
; GENERAL INFORMATION:
; APPLICANT: de Boer, Mark

COUNTRY: USA

US-08-606-293-2.
; Sequence 2, Application US/08606293
; Patent No. 5874082
; GENERAL INFORMATION:
; APPLICANT: de Boer, Mark
; TITLE OF INVENTION: Humanized Anti-CD40 Monoclonal Antibodies and
; TITLE OF INVENTION: Fragments Capable of Blocking B Cell Activation
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CHIRON CORPORATION Intellectual Property - R440
; STREET: 4560 Horton Street, P.O. Box 8097
; CITY: Emeryville
; STATE: California
; COUNTRY: United States of America
; ZIP: 94662-8097
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/606,293
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/070,158
; FILING DATE: 28-MAY-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Savereide, Paul B.
; REGISTRATION NUMBER: 36,914
; REFERENCE/DOCKET NUMBER: 27527/33157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (510) 601-2718
; TELEFAX: (510) 655-3542
; TELEX: n/a
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 113 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-606-293-2

Query Match 75.3%; Score 487; DB 2; Length 113;
Best Local Similarity 78.5%; Pred. No. 3.5e-42;
Matches 95; Conservative 9; Mismatches 9; Indels 8; Gaps 2;

1 QVQLKESGPGLVKPSQTLSTICTISGFSLSRYSVHWVRQPPGKGLWLGMMGGGNTDYN 60
||:||||| ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1 QVKLEESGPGLVAPSQLSITCTVSGFSLRSYVWVRQPPGKGLWLGMMGGGSTDYN 60

61 SALKSRLSISKDNKNOVFLKMNLSLTAADTAVYYCARKGEFYGYDGFVYWGQTLVTVS 120
||||| ||:||||| ||:||||| ||:||||| ||:||||| ||:||||| ||:||||| ||:|||||
61 SALKSRLSISKDTSKQVFLKMNLSLQTDOTAMYYCVRT-----DG-DYWGQTSVTVS 112

121 S 121
|
113 S 113

RESULT 20
US-08-652-558-38
; Sequence 38, Application US/08652558
; Patent No. 5861155
; GENERAL INFORMATION:
; APPLICANT: LIN, AUGUSTINE YEE-THARN
; TITLE OF INVENTION: HUMANIZED ANTIBODIES AND USES
; TITLE OF INVENTION: THEREOF
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BANNER & WITCOFF
; STREET: 75 STATE STREET, 23RD FLOOR
; CITY: BOSTON

; STATE: MASSACHUSETTS
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch diskette, 1.4 Mb storage
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Wordperfect 6.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/652,558
; FILING DATE: JUNE 6, 1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/IB94/00387
; FILING DATE: NOVEMBER 21, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: YANKWICH, LEON R.
; REGISTRATION NUMBER: 30,237
; REFERENCE/DOCKET NUMBER: 95,497-L
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-345-9100
; TELEFAX: 617-345-9111
; INFORMATION FOR SEQ ID NO: 38:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 120 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-652-558-38

Query Match 75.2%; Score 486.5; DB 2; Length 120;
Best Local Similarity 78.9%; Pred. No. 4.3e-42;
Matches 97; Conservative 5; Mismatches 16; Indels 5; Gaps 2;

QY 1 QVQLKESGPGLVKPSQTLSTICTISGFSLSRYSVHWVRQPPGKGLWLGMMGGGNTDYN 60
||||| ||:||||| ||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1 QVQLKESGPGLVAPSQLSITCTVSGFSLTAYGVNWRQPPGKGLWLGMMGGDNTDYN 60

QY 61 SALKSRLSISKDNKNOVFLKMNLSLTAADTAVYYCARK--GEFYGYDGFVYWGQTLVT 118
||||| ||:||||| ||:||||| ||:||||| ||:||||| ||:||||| ||:||||| ||:|||||
Db 61 SALKSRLSISKDNKSQVFLKMNLSLQTDOTARYYCARDRVATLYAMD---YWGQTSVT 117

QY 119 VSS 121
|||
Db 118 VSS 120

RESULT 21
US-08-190-199A-67
; Sequence 67, Application US/08190199A
; Patent No. 5830663
; GENERAL INFORMATION:
; APPLICANT: EMBLETON, Michael J.
; APPLICANT: GOROCHOV, Guy
; APPLICANT: JONES, Peter T.
; APPLICANT: WINTER, Gregory P.
; TITLE OF INVENTION: TREATMENT OF CELL POPULATIONS
; NUMBER OF SEQUENCES: 70
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PILLSBURY MADISON & SUTRO, L.L.P.
; STREET: 1100 New York Avenue, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3918
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Microsoft Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/190,199A
; FILING DATE: 13-JUL-1994

CLASSIFICATION: 435
PRIOR APPLICATION DATA: PCT/GB92/01483
FILING DATE: 10-AUG-1992
PRIOR APPLICATION DATA: GB 9212419.7
FILING DATE: 11-JUN-1992
PRIOR APPLICATION DATA: GB 9117352.6
FILING DATE: 10-AUG-1991
APPLICATION NUMBER: GB 9212419.7
FILING DATE: 11-JUN-1992
APPLICATION NUMBER: GB 9117352.6
FILING DATE: 10-AUG-1991
INFORMATION FOR SEQ ID NO: 67:
SEQUENCE CHARACTERISTICS:
LENGTH: 222 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-190-199A-67

Query Match 75.1%; Score 486; DB 2; Length 222;
Best Local Similarity 76.9%; Pred. No. 9.8e-42;
Matches 93; Conservative 9; Mismatches 11; Indels 8; Gaps 1;
QY 1 QVQLKESGPGLVKPSQTLTITCTISGFSLSRYSVHWVRQPPGKGLWLGMIWGGNTDYN 60
Db 1 QVQLKESGPGLVAPSQSLTITCTVSGFSLTSYGVHWVRQPPGKGLWLGVIWAGGSTNYN 60
QY 61 SALKSRLSISKDMSKQVFLKMNLSLTAAADTAVYYCARKGEFYGYDGFVYWGQGTLTVTVS 120
Db 61 SALKSRLSISKDMSKQVFLKMNLSLTAAADTAVYYCARKGEFYGYDGFVYWGQGTLTVTVS 112
QY 121 S 121
Db 113 A 113

RESULT 22
US-08-190-199A-61
Sequence 61, Application US/08190199A
Patent No. 5830663
GENERAL INFORMATION:
APPLICANT: EMBLETON, Michael J.
APPLICANT: GOROCHOV, Guy
APPLICANT: JONES, Peter T.
APPLICANT: WINTER, Gregory P.
TITLE OF INVENTION: TREATMENT OF CELL POPULATIONS
NUMBER OF SEQUENCES: 70
CORRESPONDENCE ADDRESS:
ADDRESSEE: PILLSBURY MADISON & SUTRO, L.L.P.
STREET: 1100 New York Avenue, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3918
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Microsoft Word
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/190,199A
FILING DATE: 13-JUL-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB92/01483
FILING DATE: 10-AUG-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9212419.7
FILING DATE: 11-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9117352.6
FILING DATE: 10-AUG-1991
INFORMATION FOR SEQ ID NO: 61:
SEQUENCE CHARACTERISTICS:

LENGTH: 235 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-190-199A-61
Query Match 75.1%; Score 486; DB 2; Length 235;
Best Local Similarity 76.9%; Pred. No. 1e-41;
Matches 93; Conservative 9; Mismatches 11; Indels 8; Gaps 1;
QY 1 QVQLKESGPGLVKPSQTLTITCTISGFSLSRYSVHWVRQPPGKGLWLGMIWGGNTDYN 60
Db 1 QVQLKESGPGLVAPSQSLTITCTVSGFSLTSYGVHWVRQPPGKGLWLGVIWAGGSTNYN 60
QY 61 SALKSRLSISKDMSKQVFLKMNLSLTAAADTAVYYCARKGEFYGYDGFVYWGQGTLTVTVS 120
Db 61 SALKSRLSISKDMSKQVFLKMNLSLTAAADTAVYYCARKGEFYGYDGFVYWGQGTLTVTVS 112
QY 121 S 121
Db 113 A 113
RESULT 23
US-08-621-751A-4
Sequence 4, Application US/08621751A
Patent No. 5882644
GENERAL INFORMATION:
APPLICANT: Chang, Chung N.
APPLICANT: Landolfi, Nicholas F.
APPLICANT: Martin, Ulrich
TITLE OF INVENTION: MONOCLONAL ANTIBODIES SPECIFIC FOR THE
TITLE OF INVENTION: PLATELET-DERIVED GROWTH FACTOR RECEPTOR BETA RECEPTOR AND
TITLE OF INVENTION: METHODS OF USE THEREOF
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORRISON & FOERSTER LLP
STREET: 755 PAGE MILL ROAD
CITY: PALO ALTO
STATE: CA
COUNTRY: USA
ZIP: 94304-1018
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/621,751A
FILING DATE: 22-MAR-1996
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Lehnhardt, Susan K.
REGISTRATION NUMBER: 33,943
REFERENCE/DOCKET NUMBER: 321152000100
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 813-5600
TELEFAX: (650) 494-0792
TELEX: 706141 MRSN FOERS SFO
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 137 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-621-751A-4
Query Match 74.4%; Score 481.5; DB 2; Length 137;
Best Local Similarity 76.0%; Pred. No. 1.6e-41;
Matches 92; Conservative 11; Mismatches 15; Indels 3; Gaps 1;
QY 1 QVQLKESGPGLVKPSQTLTITCTISGFSLSRYSVHWVRQPPGKGLWLGMIWGGNTDYN 60

Db 20 QVQLKESGPGLVAPSQSLITCTVSGFSLTNYAINWVRQPPQGGLWLGIIWTGGTSYN 79
QY 61 SALKSRLSISKDNQVFLKMNLSLTAADTAVYYCARKGEFYGYDGFVYWGQGLTVTS 120
Db 80 SALKSRLSISKDNQVFLKMNLSLTDGTTARYYCARTGTRGYFFD---YWGQGTTLTVS 136
QY 121 S 121
Db 137 S 137
RESULT 24
US-08-652-558-7
; Sequence 7, Application US/08652558
; Patent No. 5861155
; GENERAL INFORMATION:
; APPLICANT: LIN, AUGUSTINE YEE-THARN
; TITLE OF INVENTION: HUMANIZED ANTIBODIES AND USES
; TITLE OF INVENTION: THEREOF
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BANNER & WITCOFF
; STREET: 75 STATE STREET, 23RD FLOOR
; CITY: BOSTON
; STATE: MASSACHUSETTS
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch diskette, 1.4 Mb storage
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Wordperfect 6.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/652,558
; FILING DATE: JUNE 6, 1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/IB94/00387
; FILING DATE: NOVEMBER 21, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: YANKWICH, LEON R.
; REGISTRATION NUMBER: 30,237
; REFERENCE/DOCKET NUMBER: 95,497-L
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-345-9100
; TELEFAX: 617-345-9111
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 120 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-652-558-7

Query Match 74.1%; Score 479.5; DB 2; Length 120;
Best Local Similarity 74.0%; Pred. NO. 2.2e-41;
Matches 91; Conservative 15; Mismatches 12; Indels 5; Gaps 2;
QY 1 QVQLKESGPGLVKPSQTLSTCTISGFSLSRYSVHVVRRQPPGKGLWLGMIWGNTDYN 60
Db 1 QVQLQESGPGLVKPSQTLSTCTISGFSLSRYSVHVVRRQPPGKGLWLGMIWGNTDYN 60
QY 61 SALKSRLSISKDNQVFLKMNLSLTAADTAVYYCARK--GEFYGYDGFVYWGQGLTVT 118
Db 61 SALKSRVTMLKDTSKNQVFLRLSSVTADTAVYYCARDRTATLYAMD---YWGQGLTVT 117
QY 119 VSS 121
Db 118 VSS 120

RESULT 25
US-09-254-189-4

; Sequence 4, Application US/09254189
; Patent No. 6150792
; GENERAL INFORMATION:
; APPLICANT: Lundquist, Tomas
; TITLE OF INVENTION: Sequence Listing
; Patent No. 6150792
; FILE REFERENCE: 3526/00000
; CURRENT APPLICATION NUMBER: US/09/254,189
; CURRENT FILING DATE: 1999-03-02
; EARLIER APPLICATION NUMBER: 00/000,000
; EARLIER FILING DATE: 1998-11-24
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 120
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:humanized
; OTHER INFORMATION: monoclonal antibody
US-09-254-189-4

Query Match 74.1%; Score 479.5; DB 4; Length 120;
Best Local Similarity 74.0%; Pred. NO. 2.2e-41;
Matches 91; Conservative 15; Mismatches 12; Indels 5; Gaps 2;
QY 1 QVQLKESGPGLVKPSQTLSTCTISGFSLSRYSVHVVRRQPPGKGLWLGMIWGNTDYN 60
Db 1 QVQLQESGPGLVKPSQTLSTCTISGFSLSRYSVHVVRRQPPGKGLWLGMIWGNTDYN 60
QY 61 SALKSRLSISKDNQVFLKMNLSLTAADTAVYYCARK--GEFYGYDGFVYWGQGLTVT 118
Db 61 SALKSRVTMLKDTSKNQVFLRLSSVTADTAVYYCARDRTATLYAMD---YWGQGLTVT 117
QY 119 VSS 121
Db 118 VSS 120

Search completed: June 23, 2003, 14:04:53
Job time : 8.07979 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 23, 2003, 14:03:36 ; Search time 17.1631 Seconds
(without alignments)
762.858 Million cell updates/sec

Title: US-10-056-052A-20
Perfect score: 647
Sequence: 1 QVQLKESGPGLVKPSQTLISI.....YYGDFGVWGGTLVTVSS 121

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 417779 seqs, 108206813 residues 417779
1 number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	647	100.0	121	9	US-10-056-052-20
2	605	93.5	121	9	US-10-056-052-12
3	538.5	83.2	118	9	US-10-056-052-16
4	524.5	81.1	118	9	US-10-056-052-8
5	491.5	76.0	119	12	US-10-140-555-2
6	484.5	74.9	116	9	US-10-194-975-100
7	476	73.6	119	9	US-10-232-187-2
8	474.5	73.3	139	1	US-08-779-784-37
9	472.5	73.0	120	9	US-10-194-975-112
10	471.5	72.9	333	9	US-10-059-261-61
11	468.5	72.4	139	10	US-09-881-823-4
12	468.5	72.4	250	9	US-10-194-975-110
13	467	72.2	119	9	US-10-144-644-15
14	467	72.2	119	9	US-10-144-644-62
15	466.5	72.1	135	1	US-08-779-784-32
16	466.5	72.1	140	10	US-09-007-093-4
17	458	70.8	152	10	US-09-990-205-4
18	457.5	70.7	114	10	US-09-865-483-8
19	454.5	70.2	112	10	US-09-824-286-3
					Sequence 20, Appli
					Sequence 3, Appli
					Sequence 11, Appli
					Sequence 12, Appli
					Sequence 13, Appli
					Sequence 14, Appli
					Sequence 15, Appli
					Sequence 16, Appli
					Sequence 17, Appli
					Sequence 18, Appli
					Sequence 19, Appli
					Sequence 20, Appli
					Sequence 21, Appli
					Sequence 22, Appli
					Sequence 23, Appli
					Sequence 24, Appli
					Sequence 25, Appli
					Sequence 26, Appli
					Sequence 27, Appli
					Sequence 28, Appli
					Sequence 29, Appli
					Sequence 30, Appli
					Sequence 31, Appli
					Sequence 32, Appli
					Sequence 33, Appli
					Sequence 34, Appli
					Sequence 35, Appli
					Sequence 36, Appli
					Sequence 37, Appli
					Sequence 38, Appli
					Sequence 39, Appli
					Sequence 40, Appli
					Sequence 41, Appli
					Sequence 42, Appli
					Sequence 43, Appli
					Sequence 44, Appli
					Sequence 45, Appli
					Sequence 46, Appli
					Sequence 47, Appli
					Sequence 48, Appli
					Sequence 49, Appli
					Sequence 50, Appli
					Sequence 51, Appli
					Sequence 52, Appli
					Sequence 53, Appli
					Sequence 54, Appli
					Sequence 55, Appli
					Sequence 56, Appli
					Sequence 57, Appli
					Sequence 58, Appli
					Sequence 59, Appli
					Sequence 60, Appli
					Sequence 61, Appli
					Sequence 62, Appli
					Sequence 63, Appli
					Sequence 64, Appli
					Sequence 65, Appli
					Sequence 66, Appli
					Sequence 67, Appli
					Sequence 68, Appli
					Sequence 69, Appli
					Sequence 70, Appli
					Sequence 71, Appli
					Sequence 72, Appli
					Sequence 73, Appli
					Sequence 74, Appli
					Sequence 75, Appli
					Sequence 76, Appli
					Sequence 77, Appli
					Sequence 78, Appli
					Sequence 79, Appli
					Sequence 80, Appli
					Sequence 81, Appli
					Sequence 82, Appli
					Sequence 83, Appli
					Sequence 84, Appli
					Sequence 85, Appli
					Sequence 86, Appli
					Sequence 87, Appli
					Sequence 88, Appli
					Sequence 89, Appli
					Sequence 90, Appli
					Sequence 91, Appli
					Sequence 92, Appli
					Sequence 93, Appli
					Sequence 94, Appli
					Sequence 95, Appli
					Sequence 96, Appli
					Sequence 97, Appli
					Sequence 98, Appli
					Sequence 99, Appli
					Sequence 100, Appli
					Sequence 101, Appli
					Sequence 102, Appli
					Sequence 103, Appli
					Sequence 104, Appli
					Sequence 105, Appli
					Sequence 106, Appli
					Sequence 107, Appli
					Sequence 108, Appli
					Sequence 109, Appli
					Sequence 110, Appli
					Sequence 111, Appli
					Sequence 112, Appli
					Sequence 113, Appli
					Sequence 114, Appli
					Sequence 115, Appli
					Sequence 116, Appli
					Sequence 117, Appli
					Sequence 118, Appli
					Sequence 119, Appli
					Sequence 120, Appli
					Sequence 121, Appli
					Sequence 122, Appli
					Sequence 123, Appli
					Sequence 124, Appli
					Sequence 125, Appli
					Sequence 126, Appli
					Sequence 127, Appli
					Sequence 128, Appli
					Sequence 129, Appli
					Sequence 130, Appli
					Sequence 131, Appli
					Sequence 132, Appli
					Sequence 133, Appli
					Sequence 134, Appli
					Sequence 135, Appli
					Sequence 136, Appli
					Sequence 137, Appli
					Sequence 138, Appli
					Sequence 139, Appli
					Sequence 140, Appli
					Sequence 141, Appli
					Sequence 142, Appli
					Sequence 143, Appli
					Sequence 144, Appli
					Sequence 145, Appli
					Sequence 146, Appli
					Sequence 147, Appli
					Sequence 148, Appli
					Sequence 149, Appli
					Sequence 150, Appli
					Sequence 151, Appli
					Sequence 152, Appli
					Sequence 153, Appli
					Sequence 154, Appli
					Sequence 155, Appli
					Sequence 156, Appli
					Sequence 157, Appli
					Sequence 158, Appli
					Sequence 159, Appli
					Sequence 160, Appli
					Sequence 161, Appli
					Sequence 162, Appli
					Sequence 163, Appli
					Sequence 164, Appli
					Sequence 165, Appli
					Sequence 166, Appli
					Sequence 167, Appli
					Sequence 168, Appli
					Sequence 169, Appli
					Sequence 170, Appli
					Sequence 171, Appli
					Sequence 172, Appli
					Sequence 173, Appli
					Sequence 174, Appli
					Sequence 175, Appli
					Sequence 176, Appli
					Sequence 177, Appli
					Sequence 178, Appli
					Sequence 179, Appli
					Sequence 180, Appli
					Sequence 181, Appli
					Sequence 182, Appli
					Sequence 183, Appli
					Sequence 184, Appli
					Sequence 185, Appli
					Sequence 186, Appli
					Sequence 187, Appli
					Sequence 188, Appli
					Sequence 189, Appli
					Sequence 190, Appli
					Sequence 191, Appli
					Sequence 192, Appli
					Sequence 193, Appli
					Sequence 194, Appli
					Sequence 195, Appli
					Sequence 196, Appli
					Sequence 197, Appli
					Sequence 198, Appli
					Sequence 199, Appli
					Sequence 200, Appli

111	10	US-09-855-271-6	Sequence 6, Appl
151	9	US-09-797-941A-2	Sequence 2, Appl
474	9	US-10-270-555-3	Sequence 3, Appl
249	10	US-09-730-374-3	Sequence 3, Appl
144	9	US-09-797-941A-8	Sequence 8, Appl
116	9	US-10-112-788-7	Sequence 7, Appl
260	9	US-10-112-788-12	Sequence 12, Appl
119	9	US-10-125-687-5	Sequence 5, Appl
119	12	US-10-025-687-5	Sequence 5, Appl
119	9	US-10-144-644-19	Sequence 19, Appl
126	9	US-10-067-800-68	Sequence 68, Appl
135	10	US-09-855-271-28	Sequence 28, Appl
221	9	US-09-972-656-80	Sequence 80, Appl
121	9	US-10-067-800-60	Sequence 60, Appl
252	9	US-09-880-748-1994	Sequence 1994, Ap
251	9	US-09-880-748-990	Sequence 990, App
116	10	US-09-881-823-18	Sequence 18, Appl
115	9	US-09-144-886-59	Sequence 59, Appl
115	9	US-09-144-886-58	Sequence 58, Appl
253	9	US-09-880-748-1339	Sequence 1339, Ap
247	9	US-09-880-748-1651	Sequence 1651, Ap
252	9	US-09-880-748-1329	Sequence 1329, Ap
254	9	US-09-880-748-1578	Sequence 1578, Ap
256	9	US-09-880-748-1745	Sequence 1745, Ap
253	9	US-09-880-748-1361	Sequence 1361, Ap
250	9	US-09-880-748-1548	Sequence 1548, Ap
139	10	US-09-809-739-9	Sequence 9, Appli
256	9	US-09-880-748-1607	Sequence 1607, Ap
123	9	US-09-995-529-12	Sequence 12, Appl
253	9	US-09-880-748-1619	Sequence 1619, Ap
119	10	US-09-253-794-9	Sequence 9, Appli
119	10	US-09-253-794-11	Sequence 11, Appl
119	10	US-09-253-794-57	Sequence 57, Appl
120	9	US-09-996-288-36	Sequence 36, Appl
120	9	US-09-996-265-36	Sequence 36, Appl
450	9	US-09-996-288-220	Sequence 220, App
450	9	US-09-996-265-220	Sequence 220, App
450	10	US-09-796-848A-37	Sequence 37, Appl
119	10	US-09-253-794-8	Sequence 8, Appli
254	9	US-09-880-748-844	Sequence 844, App
255	9	US-09-880-748-841	Sequence 841, App
143	10	US-09-881-823-8	Sequence 8, Appli
253	9	US-09-880-748-954	Sequence 954, App
248	9	US-09-880-748-1360	Sequence 1360, Ap
230	9	US-09-972-656-72	Sequence 72, Appl
249	9	US-09-880-748-1321	Sequence 1321, Ap
254	9	US-09-880-748-1659	Sequence 1659, Ap
450	9	US-09-996-288-232	Sequence 232, App
450	9	US-09-996-288-234	Sequence 234, App
450	9	US-09-996-288-236	Sequence 236, App
450	9	US-09-996-288-236	Sequence 236, App
450	9	US-09-996-265-232	Sequence 232, App
450	9	US-09-996-265-234	Sequence 234, App
450	9	US-09-996-265-236	Sequence 236, App
120	9	US-10-194-975-114	Sequence 114, App
251	9	US-09-880-748-1538	Sequence 1538, Ap
251	9	US-09-880-748-1510	Sequence 1510, Ap
120	9	US-09-996-288-40	Sequence 40, Appl
120	9	US-09-996-288-44	Sequence 44, Appl
120	9	US-09-996-265-40	Sequence 40, Appl
120	9	US-09-996-265-40	Sequence 44, Appl
120	9	US-09-996-265-44	Sequence 36, Appl
120	10	US-09-796-848A-36	Sequence 957, App
249	9	US-09-880-748-957	Sequence 1401, Ap
252	9	US-09-880-748-1401	Sequence 222, App
450	9	US-09-996-288-222	Sequence 224, App
450	9	US-09-996-288-224	Sequence 222, App
450	9	US-09-996-265-222	Sequence 224, App
450	9	US-09-996-265-224	Sequence 39, Appl
450	10	US-09-796-848A-39	Sequence 41, Appl
450	10	US-09-796-848A-41	Sequence 13, Appl
463	9	US-10-153-382-13	Sequence 1602, Ap
253	9	US-09-880-748-1602	Sequence 18, Appl
120	9	US-09-771-415-18	Sequence 20, Appl
120	9	US-09-771-415-20	

93 382.5 59.1 120 9 US-09-771-415-22 Sequence 22, Appl
94 382.5 59.1 120 9 US-09-771-415-24 Sequence 24, Appl
95 382.5 59.1 120 9 US-09-771-415-26 Sequence 26, Appl
96 382.5 59.1 120 9 US-09-996-288-24 Sequence 24, Appl
97 382.5 59.1 120 9 US-10-011-931-3 Sequence 3, Appl
98 382.5 59.1 120 9 US-09-996-265-24 Sequence 24, Appl
99 382.5 59.1 229 10 US-09-974-449-37 Sequence 37, Appl
100 382.5 59.1 450 9 US-09-996-288-210 Sequence 210, App

ALIGNMENTS

RESULT 1
US-10-056-052-20
; Sequence 20, Application US/10056052
; Publication No. US20030099656A1
; GENERAL INFORMATION:
; APPLICANT: PATTI, Joseph M
; APPLICANT: HUTCHINS, Jeff T
; APPLICANT: DOMANSKI, Paul
; APPLICANT: PATEL, Pratiksha
; APPLICANT: HALL, Andrea
; TITLE OF INVENTION: MONOCLONAL ANTIBODIES TO THE CLFA PROTEIN
; FILE REFERENCE: P07069US04/BAS
; CURRENT APPLICATION NUMBER: US/10/056,052
; CURRENT FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 60/308,116
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 60/298,413
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: 60/274,611
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: 60/264,072
; PRIOR FILING DATE: 2001-01-26
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 20
; LENGTH: 121
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-10-056-052-20

Query Match 100.0%; Score 647; DB 9; Length 121;
Best Local Similarity 100.0%; Pred. No. 1.1e-47;
Matches 121; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 QVQLKESGPGLVKPSQTLISITCTISGFSLRSYSVHWVRQPPGKGLEWLGMIWGCGNTDYN 60
|||||
1 QVQLKESGPGLVKPSQTLISITCTISGFSLRSYSVHWVRQPPGKGLEWLGMIWGCGNTDYN 60

61 SALKSRLSISKDNSKNQVFLKMNSLTAADTAVYYCARKGEFYYGYDGFVYWGQGLTVTVS 120
|||||
61 SALKSRLSISKDNSKNQVFLKMNSLTAADTAVYYCARKGEFYYGYDGFVYWGQGLTVTVS 120

121 S 121
121 S 121

RESULT 2
US-10-056-052-12
; Sequence 12, Application US/10056052
; Publication No. US20030099656A1
; GENERAL INFORMATION:
; APPLICANT: PATTI, Joseph M
; APPLICANT: HUTCHINS, Jeff T
; APPLICANT: DOMANSKI, Paul
; APPLICANT: PATEL, Pratiksha
; APPLICANT: HALL, Andrea
; TITLE OF INVENTION: MONOCLONAL ANTIBODIES TO THE CLFA PROTEIN
; FILE REFERENCE: P07069US04/BAS
; CURRENT APPLICATION NUMBER: US/10/056,052

; CURRENT FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 60/308,116
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 60/298,413
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: 60/274,611
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: 60/264,072
; PRIOR FILING DATE: 2001-01-26
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 12
; LENGTH: 121
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-10-056-052-12

Query Match 93.5%; Score 605; DB 9; Length 121;
Best Local Similarity 92.6%; Pred. No. 3.7e-44;
Matches 112; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 QVQLKESGPGLVKPSQTLISITCTISGFSLRSYSVHWVRQPPGKGLEWLGMIWGCGNTDYN 60
|||||
Db 1 QVQLKESGPGLVAPSQSLISITCAISGFSLRSYSVHWVRQPPGKGLEWLGMIWGCGNTDYN 60

QY 61 SALKSRLSISKDNSKNQVFLKMNSLTAADTAVYYCARKGEFYYGYDGFVYWGQGLTVTVS 120
|||||
Db 61 SALKSRLSISKDNSKSQVFLKMNSLQTDATAMYYCARKGEFYYGYDGFVYWGQGLTVTVS 120

QY 121 S 121
Db 121 A 121

RESULT 3
US-10-056-052-16
; Sequence 16, Application US/10056052
; Publication No. US20030099656A1
; GENERAL INFORMATION:
; APPLICANT: PATTI, Joseph M
; APPLICANT: HUTCHINS, Jeff T
; APPLICANT: DOMANSKI, Paul
; APPLICANT: PATEL, Pratiksha
; APPLICANT: HALL, Andrea
; TITLE OF INVENTION: MONOCLONAL ANTIBODIES TO THE CLFA PROTEIN
; FILE REFERENCE: P07069US04/BAS
; CURRENT APPLICATION NUMBER: US/10/056,052
; CURRENT FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 60/308,116
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 60/298,413
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: 60/274,611
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: 60/264,072
; PRIOR FILING DATE: 2001-01-26
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 16
; LENGTH: 118
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-10-056-052-16

Query Match 83.2%; Score 538.5; DB 9; Length 118;
Best Local Similarity 84.3%; Pred. No. 1.5e-38;
Matches 102; Conservative 6; Mismatches 10; Indels 3; Gaps 1;

QY 1 QVQLKESGPGLVKPSQTLISITCTISGFSLRSYSVHWVRQPPGKGLEWLGMIWGCGNTDYN 60
|||||
Db 1 QVQLKESGPGLVAPSQSLISITCTVSGFSLRSYSVHWVRQPPGKGLEWLGMIWGCGNTDYN 60

QY 61 SALKSRLSISKDNSKNQVFLKMNSLTAADTAVYYCARKGEFYYGYDGFVYWGQGLTVTVS 120


```
Db      61 SALKSRLSITKDNKSKQVFLKMNSLQTDGTTAMYYCATA---YVGNWFWAYWGQGLTVTVS 117
QY      121 S 121
Db      118 A 118

RESULT 4
US-10-056-052-8
; Sequence 8, Application US/10056052
; Publication No. US20030099656A1
; GENERAL INFORMATION:
; APPLICANT: PATTI, Joseph M
; APPLICANT: HUTCHINS, Jeff T
; APPLICANT: DOMANSKI, Paul
; APPLICANT: PATEL, Pratiksha
; APPLICANT: HALL, Andrea
; TITLE OF INVENTION: MONOCLONAL ANTIBODIES TO THE CLFA PROTEIN
; FILE REFERENCE: P07069US04/BAS
; CURRENT APPLICATION NUMBER: US/10/056,052
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 60/308,116
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 60/298,413
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: 60/274,611
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: 60/264,072
; PRIOR FILING DATE: 2001-01-26
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 118
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-10-056-052-8

Query Match      81.1%; Score 524.5; DB 9; Length 118;
Best Local Similarity 81.8%; Pred. No. 2.2e-37;
Matches 99; Conservative 7; Mismatches 12; Indels 3; Gaps 1;

QY      1 QVQLKESGPGLVKPSQTLTITCTISGFSLSRYSVHWVRQPPGKGLWLGMIWGGNTDYN 60
Db      1 QVHLKESGPGLVAPSQSLTITCTVSGFSLRYNIHWVRQPPGKGLWLGMIWGGNTDYN 60
QY      61 SALKSRLSISKDNKSKNQVFLKMNSLTAADTAVYYCARKGEFYGYDGFFVYWGQGLTVTVS 120
Db      61 SALKSRLSISKDNKSKNQVFLKMNSLQTDGTTAMYYCA---SAYVGNWFWAYWGQGLTVTVS 117
QY      121 S 121
Db      118 A 118

RESULT 5
US-10-140-555-2
; Sequence 2, Application US/10140555
; Patent No. US2002012727A1
; GENERAL INFORMATION:
; APPLICANT: Julie A. Abrahamson
; APPLICANT: Stephen D. Holmes
; APPLICANT: Jeffrey R. Jackson
; TITLE OF INVENTION: RHAMM Antagonist Antibodies
; FILE REFERENCE: P50857
; CURRENT APPLICATION NUMBER: US/10/140,555
; CURRENT FILING DATE: 2002-05-07
; PRIOR APPLICATION NUMBER: US/09/443,790
; PRIOR FILING DATE: 1999-11-19
; PRIOR APPLICATION NUMBER: 60/109,041
; PRIOR FILING DATE: 1998-11-19
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSEQ for Windows Version 3.0
```

```
; SEQ ID NO 2
; LENGTH: 119
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-140-555-2

Query Match      76.0%; Score 491.5; DB 12; Length 119;
Best Local Similarity 79.3%; Pred. No. 1.4e-34;
Matches 96; Conservative 6; Mismatches 16; Indels 3; Gaps 1;

QY      1 QVQLKESGPGLVKPSQTLTITCTISGFSLSRYSVHWVRQPPGKGLWLGMIWGGNTDYN 60
Db      1 QVQLKESGPGLVAPSQSLTITCTVSGFSLTGYGVNVRQPPGKGLWLGMIWVDGGTDYN 60
QY      61 SALKSRLSISKDNKSKNQVFLKMNSLTAADTAVYYCARKGEFYGYDGFFVYWGQGLTVTVS 120
Db      61 SALKSRLSISKDNKSKQVFLKMNSLQTDGTTARYYCARGSSLL---GFAYWGQGLTVTVS 117
QY      121 S 121
Db      118 A 118

RESULT 6
US-10-194-975-100
; Sequence 100, Application US/10194975
; Publication No. US20030039649A1
; GENERAL INFORMATION:
; APPLICANT: Foote, Jefferson
; TITLE OF INVENTION: Super Humanized Antibodies
; FILE REFERENCE: 501231.01
; CURRENT APPLICATION NUMBER: US/10/194,975
; CURRENT FILING DATE: 2002-10-10
; PRIOR APPLICATION NUMBER: US 60/305,111
; PRIOR FILING DATE: 2001-07-12
; NUMBER OF SEQ ID NOS: 122
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 100
; LENGTH: 116
; TYPE: PRT
; ORGANISM: Mus sp.
US-10-194-975-100

Query Match      74.9%; Score 484.5; DB 9; Length 116;
Best Local Similarity 78.5%; Pred. No. 5.2e-34;
Matches 95; Conservative 9; Mismatches 12; Indels 5; Gaps 2;

QY      1 QVQLKESGPGLVKPSQTLTITCTISGFSLSRYSVHWVRQPPGKGLWLGMIWGGNTDYN 60
Db      1 QVQLQESGPGLVAPSQSLTITCTVSGFSLTGYGVNVRQPPGKGLWLGMIWGDGNTDYN 60
QY      61 SALKSRLSISKDNKSKNQVFLKMNSLTAADTAVYYCARKGEFYGYDGFFVYWGQGLTVTVS 120
Db      61 SALKSRLSISKDNKSKQVFLKMNSLHTDGTARYYCAREDRD---YRLD---YWGQGLTVTVS 115
QY      121 S 121
Db      116 S 116

RESULT 7
US-10-232-187-2
; Sequence 2, Application US/10232187
; Publication No. US20030092091A1
; GENERAL INFORMATION:
; APPLICANT: Abrahamson, Julie A.
; APPLICANT: Bochner, Bruce
; APPLICANT: Erickson-Miller, Connie L.
; APPLICANT: Kikly, Kristine K.
; APPLICANT: Schleimer, Robert
; APPLICANT: Nulku, Turkan E.
; TITLE OF INVENTION: Sialoadhesin Factor-2 Antibodies
; FILE REFERENCE: GH50042-1
```

```

; CURRENT APPLICATION NUMBER: US/10/232.187
; CURRENT FILING DATE: 2002-08-29
; PRIOR APPLICATION NUMBER: 60/187,595
; PRIOR FILING DATE: 2000-03-07
; PRIOR APPLICATION NUMBER: PCT/US01/07193
; PRIOR FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: 60/315,943
; PRIOR FILING DATE: 2001-08-30
; PRIOR APPLICATION NUMBER: 60/349,830
; PRIOR FILING DATE: 2002-01-18
; PRIOR APPLICATION NUMBER: 60/394,741
; PRIOR FILING DATE: 2002-07-10
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 119
; TYPE: PRT
; ORGANISM: Homo sapien
US-10-232-187-2
Query Match      73.6%; Score 476; DB 9; Length 119;
Best Local Similarity 75.6%; Pred. No. 2.8e-33;
Matches 93; Conservative 10; Mismatches 14; Indels 6; Gaps 3;
QY 1 QVQLKESGPGLVKPSQTLTISCTISGFSLSRYSVHWVRQPPGKGLEWLGMIWGGGNTDYN 60
Db 1 QVQLKESGPGLVAPSQSLTISCTVSGFSLI-YGAHWVRQPPGKGLEWLGVIWAGGSTNYN 59
QY 61 SALKSRLSISKDNSKNQVFLKMNSLTAAADTAVYYCARKGE--FYGYDGFVYWGQGLVT 118
Db 60 SALMSRLSISKDNSKSQVFLKINSLOTDDTALYYCARDGSSPYYYSME--YWGQGSTVT 116
QY 119 VSS 121
Db 117 VSS 119
RESULT 8
US-08-779-784-37
; Sequence 37, Application US/08779784
; Patent No. US20020164325A1
; GENERAL INFORMATION:
; APPLICANT: Rodriguez, Moses
; APPLICANT: Miller, David J.
; APPLICANT: Asakura, Kunihiro
; TITLE OF INVENTION: PROMOTION OF CENTRAL NERVOUS SYSTEM
; TITLE OF INVENTION: REMYELINATION USING MONOCLONAL AUTOANTIBODIES
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David A. Jackson, Esq.
; STREET: 411 Hackensack Ave, Continental Plaza, 4th
; STREET: Floor
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/779,784
; FILING DATE: 07-JAN-1997
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/692,084
; FILING DATE: 08-AUG-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/236,520
; FILING DATE: 29-APR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
```

```

; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 1199-1-001 CIPA
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-487-5800
; TELEFAX: 201-343-1684
; INFORMATION FOR SEQ ID NO: 37:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 139 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEetical: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Mus musculus
US-08-779-784-37
Query Match      73.3%; Score 474.5; DB 1; Length 139;
Best Local Similarity 74.0%; Pred. No. 4.2e-33;
Matches 91; Conservative 13; Mismatches 14; Indels 5; Gaps 3;
QY 1 QVQLKESGPGLVKPSQTLTISCTISGFSLSRYSVHWVRQPPGKGLEWLGMIWGGGNTDYN 60
Db 20 QVQLRQSGPGLVQPSQSLTISCTVSGFSLSYGVHWFQRSPGKGLEWLGVIWGGGSTDYN 79
QY 61 SALKSRLSISKDNSKNQVFLKMNSLTAAADTAVYYCAR-KGEF--YYGYDGFVYWGQGLVT 118
Db 80 AAFISRLSISKDNSKSQVFFKMNSLQANDTAIYYCARNRGRNYAMD---YWGQGSTVT 136
QY 119 VSS 121
Db 137 VSS 139
RESULT 9
US-10-194-975-112
; Sequence 112, Application US/10194975
; Publication No. US20030039649A1
; GENERAL INFORMATION:
; APPLICANT: Foote, Jefferson
; TITLE OF INVENTION: Super Humanized Antibodies
; FILE REFERENCE: 501231.01
; CURRENT APPLICATION NUMBER: US/10/194,975
; CURRENT FILING DATE: 2002-10-10
; PRIOR APPLICATION NUMBER: US 60/305,111
; PRIOR FILING DATE: 2001-07-12
; NUMBER OF SEQ ID NOS: 122
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 112
; LENGTH: 120
; TYPE: PRT
; ORGANISM: Mus sp.
US-10-194-975-112
Query Match      73.0%; Score 472.5; DB 9; Length 120;
Best Local Similarity 75.6%; Pred. No. 5.5e-33;
Matches 93; Conservative 11; Mismatches 14; Indels 5; Gaps 3;
QY 1 QVQLKESGPGLVKPSQTLTISCTISGFSLSRYSVHWVRQPPGKGLEWLGMIWGGGNTDYN 60
Db 1 EVKLQSGPGLVTPSQSLTISCTVSGFSLSDYGVHWRQSPGQGLEWLGVIWAGGSTNYN 60
QY 61 SALKSRLSISKDNSKNQVFLKMNSLTAAADTAVYYCAR-KG-EFYGYDGFVYWGQGLVT 118
Db 61 SALMSRKSISKDNSKSQVFLKMNSLQADDTAVYYCARKDGYSYYSMD---YWGQGSTVT 117
QY 119 VSS 121
Db 118 VSS 120
RESULT 10
```


COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/779,784
FILING DATE: 07-JAN-1997
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/692,084
FILING DATE: 08-AUG-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/236,520
FILING DATE: 29-APR-1994
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 1199-1-001 CIPA
TELEPHONE: 201-487-5800
TELEFAX: 201-343-1684
INFORMATION FOR SEQ ID NO: 32:
SEQUENCE CHARACTERISTICS:
LENGTH: 135 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE:
ORGANISM: Mus musculus
US-08-779-784-32

Query Match 72.1%; Score 466.5; DB 1; Length 135;
Best Local Similarity 75.2%; Pred. No. 2e-32;
Matches 91; Conservative 10; Mismatches 15; Indels 5; Gaps 2;
QY 1 QVQLKESGPGLVKPSQTLSTICTISGFSLSRYSVHWVRQPPGKGLWLGMIWGGGNTDYN 60
Db 20 QVQLKQSGPGLVQPSQSLSTICTVSGFSLTSYGVHWVRQSPGKGLWLGVIWSGGSTDYN 79
QY 61 SALKSRLSISKDNQVFLKMNLSLTAAADTAVYYCARKGEFYGYDGFVYWGQGLTVTS 120
Db 80 AAFISRLSISKDASKSQVFFKMNLSLHATDTAIYYCARD----YGSRG-DYWGQGSTVTS 134
121 S 121
135 S 135

RESULT 16
US-09-007-093-4
Sequence 4, Application US/09007093
Patent No. US20020025315A1
GENERAL INFORMATION:
APPLICANT: Anand, Naveen N
APPLICANT: Barber, Brian H
APPLICANT: Cates, George A
APPLICANT: Caterini, Judith E
APPLICANT: Klein, Michel H
TITLE OF INVENTION: CHIMERIC ANTIBODIES FOR DELIVERY OF
TITLE OF INVENTION: ANTIGENS TO SELECTED CELLS OF THE IMMUNE SYSTEM
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sim & McBurney
STREET: Suite 701, 330 University Avenue
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5G 1R7
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/007,093
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/483,576
FILING DATE: 07-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Stewart, Michael I
REGISTRATION NUMBER: 24,973
REFERENCE/DOCKET NUMBER: 1038-765
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 595-1155
TELEFAX: (416) 595-1163
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 140 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-09-007-093-4

Query Match 72.1%; Score 466.5; DB 10; Length 140;
Best Local Similarity 74.0%; Pred. No. 2e-32;
Matches 91; Conservative 12; Mismatches 15; Indels 5; Gaps 3;
QY 1 QVQLKESGPGLVKPSQTLSTICTISGFSLSRYSVHWVRQPPGKGLWLGMIWGGGNTDYN 60
Db 21 QVQLKESGPGLVAPSQSLSTICTVSGFSLTSYGVHWVRQPPGKGLWLGVIWAGGSINYN 80
QY 61 SALKSRLSISKDNQVFLKMNLSLTAAADTAVYYCARK-GEF-YYGYDGFVYWGQGLTVT 118
Db 81 SALMSRLSISKDNFKSQVFLKMSSLQTDGDTAMYYCARAYGDYVHYAMD---YWGQGSTVT 137
QY 119 VSS 121
Db 138 ASS 140

RESULT 17
US-09-990-205-4
Sequence 4, Application US/09990205
Patent No. US20020150572A1
GENERAL INFORMATION:
APPLICANT: FOON, Kenneth A.
APPLICANT: CHATTERJEE, Malaya
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE TREATMENT OF PSORIASIS
FILE REFERENCE: 304142000501
CURRENT APPLICATION NUMBER: US/09/990,205
CURRENT FILING DATE: 2001-11-20
PRIOR APPLICATION NUMBER: U.S. 09/192,838
PRIOR FILING DATE: 1998-11-16
PRIOR APPLICATION NUMBER: U.S. 60/065,774
PRIOR FILING DATE: 1997-11-17
NUMBER OF SEQ ID NOS: 5
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 4
LENGTH: 152
TYPE: PRT
ORGANISM: Mus Musculus
US-09-990-205-4

Query Match 70.8%; Score 458; DB 10; Length 152;
Best Local Similarity 72.7%; Pred. No. 1.1e-31;
Matches 88; Conservative 10; Mismatches 19; Indels 4; Gaps 1;
QY 1 QVQLKESGPGLVKPSQTLSTICTISGFSLSRYSVHWVRQPPGKGLWLGMIWGGGNTDYN 60
Db 20 QVQVKESGPGFLVPPSQSLSTICTVSGFSLTITVGVSWIRQPPGKGLWLGAIWGDGTTNYH 79

CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHYE P.C.
STREET: 1100 NORTH GLEBE ROAD
CITY: ARLINGTON
STATE: VIRGINIA
COUNTRY: U.S.A.
ZIP: 22201-4714
COMPUTER READABLE FORM:
MEDIUM TYPE: 1.44 Mb disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: MS Word
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/797,941A
FILING DATE: 05-Mar-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/849,410
FILING DATE: 30-MAY-1997
APPLICATION NUMBER: PCT/GB95/02777
FILING DATE: 28-NOV-1995
APPLICATION NUMBER: GB 9424449.8
FILING DATE: 02-DEC-1994
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 144 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 8:
US-09-797-941A-8

Query Match 67.4%; Score 436; DB 9; Length 144;
Best Local Similarity 68.6%; Pred. No. 7.7e-30;
Matches 81; Conservative 19; Mismatches 12; Indels 6; Gaps 2;

1 QVOLKESGPGLVKPSQTLSTCTISGFSLTAYGVHVIQSPGKGLEWLGVIWSGGGTDYN 60
61 SALKSRLSISKDnskQVFLKMNSLTAADTAVYYCARKGEFYGYDGFVYMGQGTTLVTVS 120
61 PAFISRLNINKDnskSQVFFKVDLSQLDDDRGIYYCVRNRNGYF-----FDSWGQGTTVTVS 115
121 S 121
116 S 116
Search completed: June 23, 2003, 14:19:14
Job time : 18.1631 secs

Query Match 67.4%; Score 436; DB 9; Length 144;
Best Local Similarity 68.6%; Pred. No. 7.7e-30;
Matches 81; Conservative 19; Mismatches 12; Indels 6; Gaps 2;
1 QVOLKESGPGLVKPSQTLSTCTISGFSLSYSVHVVRRQPPGKGLEWLGMIWGGGNTDYN 60
31 QVOLKESGPGLVKPSQTLSTCTISGFSLSYGVHVVRRQPPGKGLEWIGVWVRGGGSTDYN 90
61 SALKSRLSISKDnskQVFLKMNSLTAADTAVYYCARKGEFYGYDGFVY--WGQGTL 116
91 AAFMSRLNITKDNskNQVSLRSLSSVTAADTAVYYCAKS---MITTGFVMDSWGQGS 144

RESULT 25
US-10-112-788-7
Sequence 7, Application US/10112788
Application No. US20030077676A1
GENERAL INFORMATION:
APPLICANT: DENARDO, SALLY
APPLICANT: WINTHROP, MICHELLE
APPLICANT: DENARDO, GERALD
TITLE OF INVENTION: ANTI-MUC-1 SINGLE CHAIN ANTIBODIES FOR TUMOR TARGETING
FILE REFERENCE: 309T-000210US
CURRENT APPLICATION NUMBER: US/10/112,788
CURRENT FILING DATE: 2002-03-28
PRIOR APPLICATION NUMBER: US 60/280,721
PRIOR FILING DATE: 2001-04-30
NUMBER OF SEQ ID NOS: 19
SOFTWARE: PatentIn version 3.0
SEQ ID NO 7
LENGTH: 116
TYPE: PRT
ORGANISM: Mus musculus
US-10-112-788-7

Query Match 67.2%; Score 434.5; DB 9; Length 116;
Best Local Similarity 66.1%; Pred. No. 8.4e-30;
Matches 80; Conservative 16; Mismatches 20; Indels 5; Gaps 1;

1 QVOLKESGPGLVKPSQTLSTCTISGFSLSYSVHVVRRQPPGKGLEWLGMIWGGGNTDYN 60

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 23, 2003, 13:59:02 ; Search time 19.5231 Seconds
(without alignments)
825.861 Million cell updates/sec.

Title: US-10-056-052A-20
Perfect score: 647
Sequence: 1 QVQLKESGPGLVKPSQTLTI.....YYGYDGFVWGQGLTVTVSS 121

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database : A_Geneseq_101002: *
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT: *
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT: *
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT: *
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT: *
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT: *
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT: *
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT: *
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT: *
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT: *
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT: *
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT: *
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT: *
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT: *
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT: *
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT: *
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT: *
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT: *
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT: *
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT: *
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT: *
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT: *
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT: *
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT: *

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	523.5	80.9	118	22	AA1970
2	523.5	80.9	581	22	AA1972
3	506	78.2	119	22	AA1975
4	506	78.2	138	22	AA1976
5	502	77.6	113	18	AA1975
6	496	76.7	121	21	AA1975
7	496	76.7	476	22	AA1975
8	494.5	76.4	137	18	AA1975
9	494	76.4	183	12	AA1975
10	494	76.4	183	14	AA1975

11	492.5	76.1	116	14	AA1951
12	492	76.0	119	20	AA19218
13	492	76.0	119	23	AA19238
14	491.5	76.0	119	21	AA19215
15	491	75.9	241	22	AA19233
16	489.5	75.7	118	22	AA19173
17	489.5	75.7	137	22	AA19175
18	488.5	75.5	135	8	AA19091
19	488.5	75.5	211	15	AA195231
20	488	75.4	127	11	AA196355
21	487.5	75.3	116	22	AA197513
22	487.5	75.3	120	20	AA192432
23	487.5	75.3	239	18	AA195561
24	487.5	75.3	239	18	AA195813
25	487.5	75.3	241	13	AA191261
26	487.5	75.3	242	11	AA196483
27	487.5	75.3	242	14	AA193680
28	487.5	75.3	242	17	AA192192
29	487.5	75.3	242	17	AA199650
30	487.5	75.3	267	11	AA194841
31	487.5	75.3	272	13	AA191260
32	487.5	75.3	285	16	AA1964810
33	487.5	75.3	1221	15	AA1952699
34	487	75.3	113	18	AA192323
35	487	75.3	138	18	AA198944
36	487	75.3	138	18	AA198942
37	486.5	75.2	120	22	AA1963986
38	486	75.1	222	14	AA192843
39	486	75.1	235	14	AA1932840
40	484.5	74.9	138	13	AA191406
41	484.5	74.9	241	20	AA198988
42	484	74.8	123	18	AA197438
43	484	74.8	240	17	AA195133
44	484	74.8	651	17	AA195135
45	484	74.8	892	17	AA195139
46	484	74.8	892	17	AA195140
47	484	74.8	1020	17	AA195141
48	483.5	74.7	241	11	AA196482
49	482	74.5	223	15	AA196236
50	482	74.5	269	13	AA192569
51	481.5	74.4	137	18	AA1930277
52	479.5	74.1	120	20	AA192431
53	479.5	74.1	138	11	AA194839
54	479.5	74.1	261	21	AA1944990
55	478.5	74.0	239	14	AA193679
56	478.5	74.0	239	17	AA199649
57	478	73.9	113	13	AA191268
58	478	73.9	142	22	AA196523
59	477.5	73.8	239	17	AA192191
60	476.5	73.6	112	23	AB190483
61	475	73.4	138	17	AA191146
62	473.5	73.2	233	16	AA194817
63	471.5	72.9	116	17	AA195823
64	471.5	72.9	273	17	AA195827
65	471.5	72.9	446	17	AA195829
66	469.5	72.6	120	23	AA192846
67	469.5	72.6	120	23	AA192854
68	469.5	72.6	510	23	AA192859
69	469.5	72.6	510	23	AA192860
70	469	72.5	253	23	AA192867
71	468.5	72.4	450	21	AA194991
72	468.5	72.4	456	21	AA194992
73	468	72.3	116	11	AA197322
74	468	72.3	116	17	AA194490
75	468	72.3	116	17	AA199877
76	467	72.2	119	17	AA198492
77	467	72.2	119	19	AA194251
78	466.5	72.1	120	20	AA192429
79	466.5	72.1	140	18	AA192538
80	466.5	72.1	140	23	AA192021
81	466.5	72.1	272	18	AA196688
82	464.5	71.8	120	20	AA192430
83	464	71.7	119	19	AA192471

84 463.5 71.6 114 22 AAU02563 Anti-adipocyte mon
85 462 71.4 119 17 AAR98478 MAb 2B6 heavy chai
86 461 71.3 142 22 AAG66520 Mouse antibody 26
87 460 71.1 115 18 AAW06203 Xenograft antibody
88 459.5 71.0 120 20 AAY22433 TM27 antibody VH c
89 458.5 70.9 120 12 AAR13310 HuVhlys. Syntheti
90 458 70.8 152 20 AAY49210 MAb 1A7 heavy chai
91 458 70.8 152 20 AAY28469 Heavy chain variab
92 458 70.8 152 20 AAY21546 Monoclonal antibod
93 458 70.8 263 20 AAY28470 Vh-(Lk)-Vl of anti
94 457.5 70.7 114 23 AAE17793 Escherichia coli m
95 456.5 70.6 120 16 AAR77303 Variable heavy cha
96 456.5 70.6 133 18 AAW10543 Humanised murine a
97 455.5 70.4 107 15 AAR52043 Heavy chain variab
98 455.5 70.4 112 13 AAR26003 H-chain variable r
99 455 70.3 119 21 AAY90818 260F9 hybridoma VL
100 454.5 70.2 112 19 AAW31648 Monoclonal antibod

ALIGNMENTS

RESULT 1
AAB81970
ID AAB81970 standard; Protein; 118 AA.
XX
AC AAB81970;
XX
DT 03-JUL-2001 (first entry)
XX
DE Ganglioside GD2 specific antibody related protein SEQ ID NO: 13.
XX
KW Ganglioside; GD2; complementation determining region; CDR; antibody;
KW mouse; cancer.
XX
OS Synthetic.
XX
PN WO200123573-A1.
XX
PD 05-APR-2001.
XX
PF 29-SEP-2000; 2000WO-JP06773.
XX
PR 30-SEP-1999; 99JP-0278290.
XX
PA (KYOW) KYOWA HAKKO KOGYO KK.
XX
PI Hanai N, Shitara K, Nakamura K, Niwa R;
XX
DR WPI; 2001-266163/27.
XX
PT Human type complementation-determining domain transplanted antibody and
PT derivatives against ganglioside GD2, useful in diagnosis and therapy of
PT e.g. tumours, has low antigenicity, little side effects but potent
PT activity in cancer -
XX
PS Claim 10; Page 102-103; 123pp; Japanese.
XX
CC The present invention describes an antibody, which can react specifically
CC with ganglioside GD2, and is transplanted with a human type
CC complementation-determining domain (CDR), or its fragments. The antibody
CC and its derivatives are useful in diagnosis and therapy of tumours,
CC particularly cancer diagnosis. The present sequence is a protein
CC used in the exemplification of the invention.
XX
SQ Sequence 118 AA;

Query Match 80.9%; Score 523.5; DB 22; Length 118;
Best Local Similarity 80.2%; Pred. No. 1.5e-41;
Matches 97; Conservative 13; Mismatches 8; Indels 3; Gaps 1;

QY 1 QVQLKESGPGLVKPSQTLTITCTVSGFSLSRYSVHWVRQPPGKGLEWLGMIWGGNTDYN 60
; ||||:|||||||||||||||||:||||: ||:|||||||||||||:||||: ||:||||

Db 1 QVQLKESGPGLVKPSQTLTITCTVSGFSLASYNIIHWVRQPPGKGLEWLGVIWAGGSTNYN 60
QY 61 SALKSRLSISKDNSKNQVFLKMNLSLTAADTAVYYCARKGEFYGYDGFVYWGQGLTVTVS 120
Db 61 SALKSRLTISKDNSKNQVFLKMNLSLTAADTAVYYCAKRSD---DYSWFAYWGQGLTVTVS 117
QY 121 S 121
Db 118 S 118
RESULT 2
AAB81972
ID AAB81972 standard; Protein; 581 AA.
XX
AC AAB81972;
XX
DT 03-JUL-2001 (first entry)
XX
DE Ganglioside GD2 specific antibody related protein SEQ ID NO: 31.
XX
KW Ganglioside; GD2; complementation determining region; CDR; antibody;
KW mouse; cancer.
XX
OS Synthetic.
XX
PN WO200123573-A1.
XX
PD 05-APR-2001.
XX
PF 29-SEP-2000; 2000WO-JP06773.
XX
PR 30-SEP-1999; 99JP-0278290.
XX
PA (KYOW) KYOWA HAKKO KOGYO KK.
XX
PI Hanai N, Shitara K, Nakamura K, Niwa R;
XX
DR WPI; 2001-266163/27.
XX
PT Human type complementation-determining domain transplanted antibody and
PT derivatives against ganglioside GD2, useful in diagnosis and therapy of
PT e.g. tumours, has low antigenicity, little side effects but potent
PT activity in cancer -
XX
PS Example 3; Page 111-114; 123pp; Japanese.
XX
CC The present invention describes an antibody, which can react specifically
CC with ganglioside GD2, and is transplanted with a human type
CC complementation-determining domain (CDR), or its fragments. The antibody
CC and its derivatives are useful in diagnosis and therapy of tumours,
CC particularly cancer diagnosis. The present sequence is a protein
CC used in the exemplification of the invention.
XX
SQ Sequence 581 AA;

Query Match 80.9%; Score 523.5; DB 22; Length 581;
Best Local Similarity 80.2%; Pred. No. 8.7e-41;
Matches 97; Conservative 13; Mismatches 8; Indels 3; Gaps 1;

QY 1 QVQLKESGPGLVKPSQTLTITCTVSGFSLSRYSVHWVRQPPGKGLEWLGMIWGGNTDYN 60
Db 1 QVQLKESGPGLVKPSQTLTITCTVSGFSLASYNIIHWVRQPPGKGLEWLGVIWAGGSTNYN 60
QY 61 SALKSRLSISKDNSKNQVFLKMNLSLTAADTAVYYCARKGEFYGYDGFVYWGQGLTVTVS 120
Db 61 SALKSRLTISKDNSKNQVFLKMNLSLTAADTAVYYCAKRSD---DYSWFAYWGQGLTVTVS 117
QY 121 S 121
Db 118 S 118

XX 23-MAR-1998 (first entry)
XX Humanized 5D12 monoclonal antibody heavy chain variable region.
DE Human; B7 antigen; CD40; monoclonal antibody; B cell; growth;
XX differentiation; allergy; autoimmune disease.
KW Chimeric - Mus sp.
XX Chimeric - Homo sapiens.
OS WO9731025-A1.
XX 28-AUG-1997.
PN 21-FEB-1997; 97WO-US02858.
XX 23-FEB-1996; 96US-0606293.
XX (CHIR) CHIRON CORP.
PA De Boer M;
XX WPI; 1997-435094/40.
DR N-PSDB; AAT79192.
XX Humanised monoclonal antibody binding CD40 antigen on B cells - to
PT prevent their growth and differentiation, useful to treat or prevent
PT antibody mediated allergies and autoimmune diseases
XX Example 8; Fig 12; 64pp; English.
PS A novel humanised monoclonal antibody (MAB) has been developed which:
XX (a) binds a human CD40 antigen on the surface of a B cell, to prevent
CC its growth and differentiation; and (b) has an effective number of
CC exposed amino acids in its framework regions that are consistent with
CC amino acid residues found in the corresponding framework regions of a
CC human Ab to provide a reduced immunogenicity in humans. The present
CC sequence represents humanized 5D12 monoclonal antibody heavy chain
CC variable region, used in an example of the present invention. The MAB
CC can be used to prevent or treat antibody mediated diseases, particularly
CC immunoglobulin E (IgE) mediated allergies, systemic lupus erythematosus,
CC primary biliary cirrhosis, idiopathic thrombocytopenia purpura and
CC rheumatoid arthritis. The MAB inhibits the normal B cell response to
CC CD40 ligands at relatively low concentrations and is unlikely to cause
CC any immune response in humans.
XX Sequence 113 AA;
SQ Query Match 77.6%; Score 502; DB 18; Length 113;
Best Local Similarity 81.0%; Pred. No. 1.4e-39;
Matches 98; Conservative 7; Mismatches 8; Indels 8; Gaps 2;
QY 1 QVQLKESGPGLVKPSQTLSTICTISGFSLSRYSVHVVRRQPPGKGLWLGMIWGGGNTDYN 60
Db 1 QVQLVESGPGLVKPSQSLSTICTVSGFSLSRYSVYVVRQPPGKGLWLGMMWGGGSTDYN 60
QY 61 SALKSRLSISKDNSKNQVFLKMNLSLTAAADTAVYYCARKGEFYGYDGFVYWGQGLTVTVS 120
Db 61 SALKSRLTISKDTSKNQVFLKMNLSRAEDTAMYCYVRT-----DG-DYWGQGTTVTVS 112
QY 121 S 121
Db 113 S 113
RESULT 6
AAB10020
ID AAB10020 standard; Protein; 121 AA.
XX
AC AAB10020;
XX
DT 01-NOV-2000 (first entry)

XX H. pylori HSP60-binding antibody heavy chain protein.
DE Acid-resistant microorganism; detection; faecal; intestine; infection;
XX monoclonal antibody; heat shock protein; HSP60; heavy chain.
KW Unidentified.
XX WO200026671-A1.
XX 11-MAY-2000.
PD 29-OCT-1999; 99WO-EP08212.
PF 29-OCT-1998; 98EP-0120517.
XX 06-NOV-1998; 98EP-0120587.
PR (CONN-) CONNEX GMBH.
PA Reiter C, Cullmann G, Friedrichs U, Heppner P, Lakner M;
XX Ringels A;
XX WPI; 2000-365747/31.
DR N-PSDB; AAA40200.
XX Detecting infection by acid-fast microbes for diagnosis of Helicobacter
PT pylori, comprises reacting a faecal sample with two binding reagents for
PT antigens that survive intestinal passage
XX Disclosure; Fig 2; 84pp; German.
PS This invention describes a novel method for the detection of a mammalian
XX infection by an acid-resistant microorganism (A) by treating a faecal
CC sample with at least two different monoclonal antibodies (MAB) (or their
CC fragments or derivatives) or aptamers (collectively (I)) and detecting
CC formation of a complex (C) between (I) and the corresponding antigen of
CC (A). The first and second (I) bind to epitopes of different antigens
CC (Ag). These epitopes are present, after passage through the intestines,
CC in at least some mammals, and have either: (i) their native structure;
CC or (ii) a structure against which an antibody is produced by an animal
CC infected or immunized with (A), or its extract, lysate, derived protein
CC or fragment, or with a synthetic peptide. Practically all mammals display
CC at least one of the specified epitopes. The method is used to detect
CC infection by acid-fast bacteria, particularly of the genera Helicobacter,
CC Mycobacterium and Campylobacter, specifically H. pylori, H. hepaticus,
CC M. tuberculosis, C. jejuni and C. pylori. (I) may also be used
CC therapeutically. The method is direct and non-invasive, and provides an
CC inexpensive and easily standardizable diagnosis, despite possible
CC degradation of antigens during passage through the intestines. This
CC sequence represents the H. pylori heat shock protein, HSP60-binding
CC antibody (DMS ACC2356) heavy chain which is used to illustrate the method
CC of the invention.
XX SQ Sequence 121 AA;
Query Match 76.7%; Score 496; DB 21; Length 121;
Best Local Similarity 79.7%; Pred. No. 5.6e-39;
Matches 98; Conservative 8; Mismatches 13; Indels 4; Gaps 3;
QY 1 QVQ-LKESGPGLVKPSQTLSTICTISGFSLSRYSVHVVRRQPPGKGLWLGMIWGGGNTDY 59
Db 1 EVQLLEESGPGLVAPSQSLSTICTVSGFSLSRYSVHVVRRQPPGKGLWLGMIWGGGSTDY 60
QY 60 NSALKSRLSISKDNSKNQVFLKMNLSLTAAADTAVYYCARK-GEFYGYDGFVYWGQGLTVT 118
Db 61 NSGLKSRLSISNDNSKQVFLKMNLSLQTDTTAIYYCARNMGGRYPDY--FDYWGQGTTLT 118
QY 119 VSS 121
Db 119 VSS 121
RESULT 7

PS Disclosure; Fig 3C; 74pp; English.

XX The invention provides a monoclonal antibody (MAB) designated 1A7, which

CC elicits an anti-GD2 (tumor-associated antigen) immunological response in

CC humans. MAB 1A7 has defined light and heavy chain variable region

CC sequences. The MAB 1A7 and polypeptides can be used for eliciting an

CC anti-GD2 immune response. The polypeptides can also be used for detecting

CC or purifying anti-GD2 antibody. The products can be used for treating GD2

CC -associated diseases, e.g. melanoma, neuroblastoma, glioma, soft tissue

CC carcinoma, and small cell carcinoma. They can be used for palliating the

CC disease or for reducing the risk of recurrence. The present sequence

CC represents the heavy chain variable region consensus sequence.

XX

SQ Sequence 119 AA;

Query Match 76.0%; Score 492; DB 20; Length 119;

Best Local Similarity 78.7%; Pred. No. 1.3e-38;

Matches 96; Conservative 8; Mismatches 14; Indels 4; Gaps 2;

1 QVQLKESGPGLVKPSQTLSTCTISGFSLSRYSVHWVRQPPGKGLWLGMIWGGNTDYN 60

1 QVQLKESGPGLVAPSQSLSTCTVSGFSLTSYGVHWVRQPPGKGLWLGVIWGDGSTNYN 60

QY 61 SALKSRLSISKDNSKNQVFLKMNSLTAADTAVYYCARK-GEFYGYDGFVYWGQGLTVTV 119

Db 61 SALKSRLSISKDNSKQVFLKMNSLQTDRTARYYCAREXXXYYAMD---YWGQGSTVTV 117

QY 120 SS 121

Db 118 SS 119

RESULT 13

AAU72838

ID AAU72838 standard; Peptide; 119 AA.

XX

AC AAU72838;

XX

DT 26-FEB-2002 (first entry)

XX

DE Anti-NKG2D hybridoma 6H7E7 variable heavy chain CDR1.

XX

KW Human; NKG2D; NKG2D receptor complex; cancer; infectious disease; tumour;

KW autoimmune disease; head; neck; stomach; oesophagus; colon; liver; ovary;

KW intrahepatic bile duct; pancreas; lung; larynx; breast; uterus; cervix;

KW prostate; kidney; testis; thyroid; bladder; brain; melanoma; myeloma; Fv;

KW sarcoma; leukaemia; lymphoma; virus; bacterium; fungus; protozoan; DAP10;

KW helminth; cytostatic; antimicrobial; immunomodulatory; 11B2D10; 6H7E7;

KW 8G7C10; 6E5A7; 11B2D10x4-7; 8G7C10x4-7; 6E5A7x4-7; P4-2; P4-3; P4-14;

KW P4-15; P5-2; P5-3; P5-9; P5-10; P5-11; P5-23; 3B10xP4-3; 3B10xP4-14;

KW P53 tetramerisation domain; 3B10xP5-2; 3B10xP5-23.

OS Homo sapiens.

XX

XX WO200171005-A2.

PN

XX

PD 27-SEP-2001.

XX

XX 26-MAR-2001; 2001WO-EP03414.

PF

XX

PR 24-MAR-2000; 2000EP-0106467.

XX

XX (KUFE/) KUFR P.

PA

XX Kufer P, Riethmueller G, Lutterbuese R, Borschert K, Kischel R;

PI Mayer M, Hofmeister R;

XX

XX WPI; 2002-055119/07.

DR

XX

PT Multifunctional polypeptides comprising binding sites that specifically

PT recognise extracellular groups of the NKG2D receptor complex and

PT domains which function as receptors or ligands, useful for treating

PT cancers and infectious diseases -

XX Example 3; Fig 16; 114pp; English.

PS

XX The invention relates to a multifunctional polypeptide comprising a

CC domain with a binding site that specifically recognises an extracellular

CC group of the NKG2D receptor complex and a second domain which functions

CC as a receptor or ligand. The polypeptide and its associated

CC polynucleotide are used for the preparation of a pharmaceutical

CC composition for the treatment of cancer, infections and/or autoimmune

CC conditions. The cancer may be a tumour of the head and neck, stomach,

CC oesophagus, colon, liver, intrahepatic bile ducts, pancreas, lung,

CC larynx, breast, ovary, uterus, cervix, prostate, kidney, testis, thyroid,

CC bladder or brain, or a melanoma, myeloma, sarcoma, leukaemia or lymphoma.

CC The infectious diseases can be caused by viruses, bacteria, fungi,

CC protozoa or helminths. The autoimmune diseases include multiple

CC sclerosis, Grave's disease, ankylosing spondylitis, acute anterior

CC uveitis, Goodpasture's syndrome, myasthenia gravis, insulin-dependent

CC diabetes mellitus, rheumatoid arthritis, pemphigus vulgaris and

CC autoimmune hepatitis. Sequences AAU72820-AAU72875 represent the NKG2D

CC receptor and the polypeptides of the invention.

XX

SQ Sequence 119 AA;

Query Match 76.0%; Score 492; DB 23; Length 119;

Best Local Similarity 77.7%; Pred. No. 1.3e-38;

Matches 94; Conservative 10; Mismatches 15; Indels 2; Gaps 1;

1 QVQLKESGPGLVKPSQTLSTCTISGFSLSRYSVHWVRQPPGKGLWLGMIWGGNTDYN 60

1 QVQLQESGPGLVAPSQSLSTCTVSGFSLTSYGVHWIRPPGKGLWLGVIWAGGSTNYN 60

QY 61 SALKSRLSISKDNSKNQVFLKMNSLTAADTAVYYCARKGEFYGYDGFVYWGQGLTVTVS 120

Db 61 SALKSRLSISKDNSKQVFLKMNSLQIDDTAMYICARGG--YEGAAWFGYWGQGTTVTVS 118

QY 121 S 121

Db 119 S 119

RESULT 14

AAU94215

ID AAU94215 standard; Protein; 119 AA.

XX

AC AAU94215;

XX

DT 08-AUG-2000 (first entry)

XX

DE Murine consensus 16E10 heavy chain variable region.

XX

KW Antibody; RHAMM; receptor for hyaluronic acid mediated motility;

KW ras-dependent proliferation; leukaemia; cancer; lymphoma;

KW inflammatory disease; proliferative disease; psoriasis;

KW inflammatory bowel disease; rheumatoid arthritis;

KW proliferative cardiovascular disease; restenosis;

KW proliferative ocular disorder; diabetic retinopathy; haemangioma;

KW benign hyperproliferative disease; tumour formation; heavy chain;

KW variable region; 16E10.

XX

OS Mus musculus.

XX

XX WO200029447-A1.

PN

XX

PD 25-MAY-2000.

XX

PF 19-NOV-1999; 99WO-US27565.

XX

PR 19-NOV-1998; 98US-0109041.

PR 14-JUL-1999; 99US-0143692.

XX

PA (SMIK) SMITHKLINE BEECHAM CORP.

PA (SMIK) SMITHKLINE BEECHAM PLC.

XX

XX (WINT/) WINTER G P.
XX Winter GP;
PI WPI; 1987-272811/39.
XX N-PSDB; AAN72111.
DR Recombinant altered antibodies - having complementarity
XX determining regions replaced with those from antibody of
PT different specificity
PT Example; Fig 7; 4lpp; English.
XX The method of the invention is used for "humanising" non-human
CC monoclonal antibodies (MABs) eg CDRs from mouse MAB can be
CC partially or totally grafted into the framework regions of a human
CC MAB, which is then produced in quantity by a suitable cell line.
CC CDS= complementarity determining regions. The antilysozyme antibody
CC D1.3 (H2K2) has two heavy chains of the mouse IgG1 class (H) and
CC two kappa light chains (K).

SQ Sequence 135 AA;
Query Match 75.5%; Score 488.5; DB 8; Length 135;
Best Local Similarity 79.3%; Pred. No. 3.2e-38;
Matches 96; Conservative 8; Mismatches 12; Indels 5; Gaps 2;

QY 1 QVQLKESGPGLVKPSQTLSTCTISGFSLSRYSVHWVRQPPGKGLWLGMIWGGNTDYN 60
|||||
Db 20 QVQLKESGPGLVAPSQSLSTCTVSGFSLTGYGVNVRQPPGKGLWLGMIWGGNTDYN 79
61 SALKSRLSISKNSKQVFLKMNLSLTAAADTAVYYCARKGEFYYGYDGFVYWGQGLTVTS 120
|||||
Db 80 SALKSRLSISKNSKQVFLKMNLSLHTDDTARYYCARERD--YRLD--YWGQGTTLTVS 134
121 S 121
135 S 135

RESULT 19
AAR56231
ID AAR56231 standard; Protein; 211 AA.
XX AAR56231;
XX 04-JAN-1995 (first entry)
XX 161-107/113-60/h1-116 D1.3 V-min intact heavy chain.

XX McPC603; V-min; synthetic gene; D1.3; antibody; immunotoxins; tumour;
KW intracellular expression; E. coli; protease deficient; treatment;
KW ligand-binding variable domain; framework region; beta-barrel;
KW beta-sheet; linker; complementarity determining regions; radioimaging;
KW produg therapy; diagnosis.
XX Synthetic.
OS Key Location/Qualifiers
XX Region 1..47 /note= "161-107"
FT Region 48..95 /note= "113-160"
FT Region 96..211 /note= "h1-h116"
XX WO9412625-A.
XX 09-JUN-1994.
XX 19-NOV-1993; 93WO-GB02375.
XX

PR 23-NOV-1992; 92GB-0024588.
PR 31-DEC-1992; 92GB-0027189.
XX (ZENE) ZENECA LTD.
XX Slater AM, Timms D;
PI WPI; 1994-200254/24.
XX New ligand variable binding domains of reduced size - contain
PT complementarity determining regions and framework region to
PT provide correct orientation, also related DNA, expression vectors
PT etc., useful for radio-imaging and tumour treatment
XX Disclosure; Fig 42; 187pp; English.
PS The sequences given in AAR56229-36 show various V-min constructs. V-min
XX is a ligand-binding variable domain which comprises a framework region
CC of a cyclically permuted central beta-barrel, outer beta-sheet
CC segments, and linker segments, and complementarity determining regions.
CC V-min, when conjugated, are useful in radioimaging, as immunotoxins
CC and in antibody directed enzyme produg therapy, ie. for diagnosis or
CC treatment of tumours. Compared to ordinary antibodies, V-min are
CC significantly smaller, as few as 160 amino acids, and so may show
CC improved tissue penetration and faster clearance from the body. Small
CC size should facilitate conjugation and oligomerisation. V-min can be
CC expressed as a continuous polypeptide from a single gene.

XX SQ Sequence 211 AA;
Query Match 75.5%; Score 488.5; DB 15; Length 211;
Best Local Similarity 79.3%; Pred. No. 5.3e-38;
Matches 96; Conservative 8; Mismatches 12; Indels 5; Gaps 2;

QY 1 QVQLKESGPGLVKPSQTLSTCTISGFSLSRYSVHWVRQPPGKGLWLGMIWGGNTDYN 60
|||||
Db 96 QVQLKESGPGLVAPSQSLSTCTVSGFSLTGYGVNVRQPPGKGLWLGMIWGGNTDYN 155
61 SALKSRLSISKNSKQVFLKMNLSLTAAADTAVYYCARKGEFYYGYDGFVYWGQGLTVTS 120
|||||
Db 156 SALKSRLSISKNSKQVFLKMNLSLHTDDTARYYCARERD--YRLD--YWGQGTTLTVS 210
121 S 121
211 S 211

RESULT 20
AAR06355
ID AAR06355 standard; peptide; 127 AA.
XX AAR06355;
XX 25-SEP-1990 (first entry)
XX Peptide corresponding to PAC1 murine monoclonal antibody.
DE Fibrinogen binding; PAC1 murine; platelet aggregation;
KW thrombosis; cancer metastasis; glycoprotein IIb-IIIa complex;
KW heavy chain; hypervariable region.
XX Synthetic.
OS Key Location/Qualifiers
XX Peptide 1..103 /label=(X1)n
FT /note="n=0 or 1"
FT 96..103 /label=(X1)n
FT /note="n=0 or 1"
FT 107..127 /label=(X2)m
FT /note="m=0 or 1"

Peptide 107..116
/label=(X2)m
/note="M=0 or 1"

EP368486-A.

16-MAY-1990.

17-OCT-1989; 89EP-0310620.

10-NOV-1988; 88US-0269583.

(MERI) MERCK AND CO INC.
(UYPE-) UNIVERSITY OF PENNSYLVANIA.

Shattil SJ, Taub RA, Friedman PA;
WPI; 1990-149427/20.

Peptide(s) corresponding to PAC1 murine monoclonal antibody -
used for inhibiting fibrinogen binding, platelet aggregation,
thrombosis and/or cancer metastasis.

Claim 8; Page 11; 17pp; English.

The peptide has the general formula: R1-(X1)n-RYD-(X2)m-R2 where R1 is a
(un)protected N-terminal gp and R2 is a (un)protected C-terminal gp.
The peptide shown is the longest given but truncated versions esp. as
indicated in the feature table are also used. The peptides inhibit the
binding of fibrinogen to blood platelets, platelet aggregation and/or
binding of cdfs. or substances, partic. proteins, to the glycoprotein
IIb-IIIa complex found in blood platelet membranes. They are therefore
useful for preventing thrombosis and cancer metastasis. The -RYD-
sequence is present in one of the heavy chain hypervariable regions of
PAC1 murine monoclonal antibody.

Sequence 127 AA;

Query Match 75.4%; Score 488; DB 11; Length 127;
Best Local Similarity 72.4%; Pred. No. 3.3e-38;
Matches 92; Conservative 11; Mismatches 18; Indels 6; Gaps 1;

1 QVQLKESGPGLVKPSQTLSTITCTISGFSLSRYSVHWVRQPPGKGLWLGMIWGGNTDYN 60
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1 QVQLKQSGPGLVQPSQSLSTICTVSGFSLTSYGVHWVRLSPGKGLWLGVIWGGSTDYN 60

61 SALKSRLSISKDNSKNQVFLKMNLSLTAADTAVYYCARKGEFYGYDG-----FVYWGQ 114
:| |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
61 AAFISRLSISKDNSKSKQVFFKMNLSQANDTGIYYCARRSPSYRYDGAGPYAYMDYWGQ 120

115 TLVTVSS 121
| |||||
121 TSVTVSS 127

RESULT 21
AAU07513
ID AAU07513 standard; Peptide; 116 AA.
XX
AC AAU07513;
XX
DT 24-OCT-2001 (first entry)
XX

Antibody scFv8 heavy chain variable region, mutant D1.3.
XX
KW Antimicrobial; antiviral; cytostatic; immunomodulatory;
KW antibody; gene therapy; HIV; human immunodeficiency virus; tumour;
KW metabolic disorder; immune disorder; auto-immune disorder; lysozyme;
KW antibody heavy chain variable region; scFv8; mutant D1.3; mutein.
OS Synthetic.
XX
FH Key Location/Qualifiers

Misc-difference 9..10 /note= "Wild-type Gly-Asp substituted by Pro-Gly"
FT
FT
Misc-difference 13 /note= "Wild-type Gln substituted by Ala"
FT
FT
Misc-difference 15..16 /note= "Wild-type Gly-Gly substituted by Ser-Gln"
FT
FT
Misc-difference 19..21 /note= "Wild-type Lys-Leu-Ser substituted by
FT Ser-Ile-Thr"
FT
Misc-difference 23..24 /note= "Wild-type Ala-Ala substituted by Thr-Val"
FT
FT
Misc-difference 28..31 /note= "Wild-type Thr-Phe-Ser-Ser substituted by
FT Ser-Leu-Thr-Gly"
FT
Misc-difference 34..35 /note= "Wild-type Met-Ser substituted by Val-Asn"
FT
FT
Misc-difference 40 /note= "Wild-type Thr substituted by Pro"
FT
FT
Misc-difference 42 /note= "Wild-type Asp substituted by Gly"
FT
FT
Misc-difference 44 /note= "Wild-type Arg substituted by Gly"
FT
FT
Misc-difference 47..50 /note= "Wild-type Leu-Val-Ala-Thr substituted by
FT Trp-Leu-Gly-Met"
FT
Misc-difference 52..54 /note= "Wil-type Asn-Ser-Asn substituted by
FT Trp-Gly-Asp"
FT
Misc-difference 56 /note= "Wild-type Ser substituted by Asn"
FT
FT
Misc-difference 58 /note= "Wild-type Phe substituted by Asp"
FT
FT
Misc-difference 60..63 /note= "Wild-type Pro-Asp-Ser-Val substituted by
FT Asn-Ser-Ala-Leu"
FT
Misc-difference 65 /note= "Wild-type Gly substituted by Ser"
FT
FT
Misc-difference 67..68 /note= "Wild-type Phe-Thr substituted by Leu-Ser"
FT
FT
Misc-difference 71 /note= "Wild-type Arg substituted by Lys"
FT
FT
Misc-difference 74 /note= "Wild-type Ala substituted by Ser"
FT
FT
Misc-difference 76..79 /note= "Wild-type Asn-Thr-Leu-Tyr substituted by
FT Ser-Gln-Val-Phe"
FT
Misc-difference 81 /note= "Wild-type Gln substituted by Lys"
FT
FT
Misc-difference 83 /note= "Wild-type Ser substituted by Asn"
FT
FT
Misc-difference 86..88 /note= "Wild-type Lys-Ser-Glu substituted by
FT His-Thr-Asp"
FT
Misc-difference 92 /note= "Wild-type Met substituted by Arg"
FT
FT
Misc-difference 98 /note= "Wild-type Arg substituted by Glu"
FT
FT
Misc-difference 100 /note= "Wild-type Asn substituted by Asp"
FT
FT
Misc-difference 102..103 /note= "Wild-type Pro-Tyr-Tyr-Tyr-Gly-Ser-Arg-Tyr-Phe
FT substituted by Arg-Leu"
FT
XX
PN WO200149713-A2.
XX
PD 12-JUL-2001.
XX
PF 29-DEC-2000; 2000WO-IT00554.
XX
PR 30-DEC-1999; 99IT-RM00803.
XX
PA (CNEN) ENEA ENTE NUOVE TECNOLOGIE ENERGIA.
PA (CONS-) SOC CONSORTILE METAPONTUM AGROBIOS SRL.

XX Benvenuto E, Franconi R, Desiderio A, Tavliadoraki P;
PI WPI; 2001-502555/55.
DR
XX Peptides which are able to confer stability and solubility to an
PT antibody comprising these peptides, useful for treating pathologies
PT (e.g. tumour) associated with accumulation of a molecule inside or
PT outside a human, or animal cell -
XX
PS Disclosure; Fig 4; 109pp; English.
XX
CC The invention relates to peptides which are able to confer stability and
CC solubility to an antibody comprising these peptides. The peptides
CC are especially H-FR1, H-FR2, HF-R3, HFR4, L-FR1, L-FR2, L-FR3 or L-FR4
CC present within a variable region of an antibody which makes the antibody
CC soluble and stable in cytoplasm. Peptides having the sequences of HFR1 to
CC H-FR4 are present within the variable region of the heavy chain of an
CC antibody, covalently linked to the H-CDR1, H-CDR2, H-CDR3 in the order
CC (H-FR1)-(H-CDR1)-(H-FR2)-(H-CDR2)-(H-FR3)-(H-CDR3)-(H-CDR4) and
CC peptides having the sequences of L-FR1 to L-FR4 are present within
CC the variable region of the light chain of an antibody, covalently
CC linked to the L-CDR1, L-CDR2, L-CDR3 in the order (L-FR1)-(L-CDR1)-(
CC L-FR2)-(L-CDR2)-(L-FR3)-(L-CDR3)-(L-CDR4). The antibodies and
CC polynucleotides (e.g. by gene therapy) are useful for the manufacture
CC of a medicament for the treatment of pathologies associated with
CC accumulation of a molecule inside or outside a human, animal cell
CC or plant cell. The pathologies are infectious (e.g. viral infections such
CC as HIV, human immunodeficiency virus, infections), tumour, metabolic and
CC immune (especially auto-immune) pathologies. The present sequence
CC represents the antibody scfv8 heavy chain variable region, mutant D1.3.
XX
SQ Sequence 116 AA;
Query Match 75.3%; Score 487.5; DB 22; Length 116;
Best Local Similarity 79.3%; Pred. No. 3.3e-38;
Matches 96; Conservative 8; Mismatches 12; Indels 5; Gaps 2;
QY 1 QVQLKESGPGLVKPSQTLSTICTISGFSLSRYSVHWVRQPPGKGLEWLGMIWGGNTDYN 60
Db 1 QVQLQESGPGLVAPSQSLSTICTVSGFSLTGYGVMVRQPPGKGLEWLGMIWGGNTDYN 60
QY 61 SALKSRLSISKDNSKNQVFLKMNSLTAADTAVYYCARKGEFFYYGYDGFVYWGQTLTVTS 120
Db 61 SALKSRLSISKDNSKSQVFLKMNSLHTDDTARYYCARERD--YRLD---YWGQGTTVTVS 115
QY 121 S 121
116 S 116
RESULT 22
AAAY22432
ID AAY22432 standard; protein: 120 AA.
XX
AC AAY22432;
XX
DT 28-SEP-1999 (first entry)
XX
DE TM27 antibody VH chain mutant V67L, T68S, M69I, L70S, T73N.
KW IgG antibody; TM27; variable chain; immunoglobulin G; medical therapy;
KW autoimmune disease; multiple sclerosis; antibody storage; VH; mutein;
KW heavy chain.
XX Homo sapiens.
OS
XX
FH Key Location/Qualifiers
FT Misc-difference 67
FT /label= V67L
FT Misc-difference 68
FT /label= T68S
FT Misc-difference 69

FT Misc-difference /label= M69I
FT 70
FT /label= L70S
FT Misc-difference 73
FT /label= T73N
XX
PN WO9937329-A1.
XX
PD 29-JUL-1999.
XX
XX 15-JAN-1999; 99WO-SE000049.
PF
XX 09-MAR-1998; 98SE-0000766.
PR 22-JAN-1998; 98SE-0000170.
XX
PA (ASTR) ASTRA AB.
XX
PI Flink O, Petren S;
XX
DR WPI; 1999-458611/38.
XX
PT Isotonic pharmaceutical antibody formulations comprising a citrate
PT buffer, have improved storage
XX
PS Claim 12; Page 23-24; 30pp; English.
XX
CC This sequence is a mutant of the variable heavy (VH) chain of the
CC antibody TM27. TM27 is an immunoglobulin G (IgG) antibody, contained in
CC the isotonic pharmaceutical formulation of the invention, along with a
CC citrate buffer at a physiologically acceptable pH. The formulations are
CC useful in medical therapy, especially for treatment of autoimmune
CC disease, and particularly, therapy of multiple sclerosis (using the TM27
CC antibody heavy or light chains). The formulation is useful for improving
CC the storage of an antibody. The antibody formulations are simple and
CC effective. The formulations are stable and have improved storage
CC properties. The formulation is simpler than prior art formulations.
CC Existing antibody formulations require the use both of a stabiliser and a
CC buffer. The present invention formulations are stabilized only by citrate
CC buffer in a saline solution at a physiologically preferable pH. The
CC avoidance of low pH prevents undesirable reaction as the site of
CC injection. The formulation does not use ovalbumin for stabilisation,
CC hence avoiding an allergic response to ovalbumin. Also the formulation
CC does not require lyophilisation which is an expensive process and also
CC requires the formulation to be resuspended prior to administration.
XX
SQ Sequence 120 AA;
Query Match 75.3%; Score 487.5; DB 20; Length 120;
Best Local Similarity 76.4%; Pred. NO. 3.5e-38;
Matches 94; Conservative 12; Mismatches 12; Indels 5; Gaps 2;
QY 1 QVQLKESGPGLVKPSQTLSTICTISGFSLSRYSVHWVRQPPGKGLEWLGMIWGGNTDYN 60
Db 1 QVQLQESGPGLVKPSQTLSTCTVSGFSLTAYGVNVRQPPGKLEWLGMIWGGNTDYN 60
QY 61 SALKSRLSISKDNSKNQVFLKMNSLTAADTAVYYCARK--GEFYYGYDGFVYWGQTLVT 118
Db 61 SALKSRLSISKDNSKNQFSLRLSSVTAADTAVYYCARDRTATLYAMD---YWGQGLVT 117
QY 119 VSS 121
Db 118 VSS 120
RESULT 23
AAW35561
ID AAW35561 standard; Protein: 239 AA.
XX
AC AAW35561;
XX
DT 24-APR-1998 (first entry)
XX
DE Ecor1-HindIII insert of .pur.4124 protein sequence.

Qy	121 S 121
Db	239 S 239

RESULT 25
AAR21261
ID AAR21261 standard; Protein: 241 AA.

AAR21261;

DT 21-MAY-1992 (first entry)

VHD1.3-HuCH1 region of Fab D1.3 in pUC19.

KW Fd; bacteriophage; gene III; filamentous; phagemid; capsid; coat;
KW pilus; g3p; binding; adsorption; gene VIII; diverse repertoire;
KW specific binding pairs; replicable genetic display package; Ig;
KW immunoglobulin; antibody; heavy chain; light chain; variable.

Synthetic.

	Key Peptide	Location/Qualifiers
FH		1..22
FT		/label= PelB_leader
FT	Region	23..138
FT		/label= VHD1.3
FT	Peptide	139..241
FT		/label= HuCH1

WO9201047-A.

23-JAN-1992.

10-JUL-1991; 91WO-GB01134.

15-MAY-1991; 91GB-0010549.

19-OCT-1990: 90GB-0022845.

12 NOV 1990, 30GB-0024303,
06-MAR-1991: .91GB-0004744

(CAMB-1) CAMBERTON ANTTRONV

(MEDI-) MED RES COUNCIL.

McCafferty J, Pope AR, John

Winter GP, Bonnet TP;

WPI; 1992-056862/07.

recombinant host cells with

THE CURIOUS

expresses a Fab-like fragment

CC to gene III and expressed on the surface of the pab.
CC See also AAR21260-307, 309-311; AAR22450, 565-581.

Sequence 241 AA;

Query Match 75.3%; Score 487.5; DB 13; Length 241;
Best Local Similarity 79.3%; Pred. No. 7.6e-38;
Matches 96; Conservative 8; Mismatches 12; Indels 5;

Qy	1 QVQLKESGPGLVKPSQTLSITCTISGFSLRSYSVHWVRQPPGKGLEWLGMIWGSGNTDYN 60 : : : : : : : : :
Db	23 OVLOESGPGVLAPSOSSLTCTVSFSGITGVGNWVRPPGKGLEWLGMIWGSGNTDYN 82 : : : : : : : : :

QY	61	SALKSRLSISKDN SKNQVFLKMNSLTAADTAVYYCAR KGEFYYGYDGFVYGQCTLT VVS.	120
Db	83	SALKSRLSISKDN SKSOVLFKMNSIHTDDTARYYCAREP---YRLD---YMGQCTTVTS	137

Ov 121 S 121

Db 1385-138

Search completed: June 23, 2003, 14:01:42
Job time : 22.5231 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
OM protein - protein search, using sw model
Run on: June 23, 2003, 13:59:02 ; Search time 18.0709 Seconds
(without alignments)
825.861 Million cell updates/sec

Title: US-10-056-052A-18
Perfect score: 583
Sequence: 1 DIVMTQSPDSLAVSLGERVT.....YCHQYLSSYTFGGGTKLEIK 112

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5
Searched: 908470 seqs, 133250620 residues

al number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database : A_Geneseq_101002:*
1: /SIDS2/gcgcdata/geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS2/gcgcdata/geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS2/gcgcdata/geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS2/gcgcdata/geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS2/gcgcdata/geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS2/gcgcdata/geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS2/gcgcdata/geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS2/gcgcdata/geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS2/gcgcdata/geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS2/gcgcdata/geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS2/gcgcdata/geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS2/gcgcdata/geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS2/gcgcdata/geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS2/gcgcdata/geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS2/gcgcdata/geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS2/gcgcdata/geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS2/gcgcdata/geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS2/gcgcdata/geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS2/gcgcdata/geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS2/gcgcdata/geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS2/gcgcdata/geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS2/gcgcdata/geneseq/geneseqp-embl/AA2001.DAT:*
23: /SIDS2/gcgcdata/geneseq/geneseqp-embl/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	551	94.5	112	15	AA154933
2	540	92.6	263	20	AA190226
3	540	92.6	268	20	AA190228
4	540	92.6	268	20	AA190222
5	540	92.6	273	20	AA190224
6	540	92.6	556	20	AA190218
7	540	92.6	580	20	AA190217
8	527	90.4	113	17	AA192215
9	527	90.4	113	18	AA192695
10	526	90.2	113	22	AA162769
					MAB 022 VK chain.
					Anti-B7.2 monospec
					Anti-B7.1/anti-B7.
					Anti-B7.2 monospec
					Anti-B7.1/anti-B7.
					Bispecific tetra
					Bispecific tetra
					LL2 MAB VK region.
					Variable kappa cha
					Human HIV-1 monoc

11	525.5	90.1	113	15	AA150313	Humanised light ch
12	524.5	90.0	120	22	AA165565	Amino acid sequenc
13	521	89.4	112	13	AA128807	5A8 VL. Synthetic
14	520.5	89.3	113	15	AA150314	Humanised light ch
15	520.5	89.3	155	18	AA132483	Kappa light chain
16	520.5	89.3	155	20	AA106912	Human variable kap.
17	520.5	89.3	342	18	AA132482	Growth factor TLHL
18	520.5	89.3	342	20	AA106909	TLHL amino acid se
19	520.5	89.3	495	18	AA132480	Growth factor CATA
20	520.5	89.3	495	20	AA106908	CATAB-TEV aminoaci
21	516.5	88.6	137	21	AA195243	Humanised antibody
22	515	88.3	241	13	AA128809	Vector pMDR1007.
23	513	88.0	259	21	AA109779	Antiviral scFv-ant
24	512.5	87.9	133	9	AA180894	y region of L chain
25	512.5	87.9	134	20	AA150690	Human Hum4 VL Clai
26	512.5	87.9	171	14	AA138320	Sequence of signal
27	512.5	87.9	171	20	AA157184	Amino acid sequenc
28	512.5	87.9	171	20	AA150693	Plasmid pATDFLAG F
29	512.5	87.9	240	22	AA100815	Hum4 Immunoglobul
30	512.5	87.9	251	22	AA174793	Single chain antib
31	512.5	87.9	258	22	AA174794	Single chain antib
32	512.5	87.9	274	14	AA138319	Sequence of Hum4 V
33	512.5	87.9	274	20	AA157183	Amino acid sequenc
34	512.5	87.9	274	20	AA150692	Human Hum4L-CC49 V
35	512.5	87.9	284	14	AA138321	Sequence of pSC49F
36	512.5	87.9	284	20	AA157185	Amino acid sequenc
37	512.5	87.9	284	20	AA150694	Plasmid pSC49FLAG
38	511	87.7	114	15	AA145605	Monoclonal antibod
39	510.5	87.6	113	21	AA190819	260F9 hybridoma VL
40	510.5	87.6	119	23	AA107172	ebvHigM MSI19p10 l
41	510	87.5	141	20	AA124374	Human monoclonal a
42	508.5	87.2	137	18	AA121653	Mouse MAB 15 light
43	508.5	87.2	137	18	AA121655	Chimeric MAB 15 PC
44	507.5	87.0	113	21	AA118881	Amino acid sequenc
45	507.5	87.0	114	14	AA130144	MAB GAH variable r
46	507.5	87.0	114	23	AA177333	Human Len kappa li
47	507.5	87.0	133	15	AA148633	Sequence of Hum4VL
48	507.5	87.0	264	23	AA143142	Human ovarian anti
49	506.5	86.9	260	20	AA157182	Amino acid sequenc
50	506.5	86.9	260	20	AA150691	Human SCFV1 protei
51	506	86.8	113	17	AA192217	Humanised LL2 MAB
52	505.5	86.7	111	15	AA152059	Light chain variab
53	505.5	86.7	113	21	AA118857	Amino acid sequenc
54	505.5	86.7	113	21	AA118873	Anti-CA125 bifunct
55	505.5	86.7	134	22	AA174792	Sequence of plasml
56	505.5	86.7	285	15	AA148638	MAB NFS2 light cha
57	504.5	86.5	113	15	AA150322	Sequence of the li
58	504.5	86.5	113	20	AA150145	Antibody F19 human
59	503.5	86.4	113	20	AA150145	Human Ab light cha
60	503.5	86.4	115	18	AA127546	Amino acid sequenc
61	502.5	86.2	113	21	AA118855	Amino acid sequenc
62	502.5	86.2	113	21	AA118871	Amino acid sequenc
63	502.5	86.2	113	21	AA118877	Amino acid sequenc
64	502.5	86.2	113	21	AA118879	Immunoglobulin kap
65	502.5	86.2	135	21	AA103714	Monoclonal antibod
66	502.5	86.2	139	17	AA199469	Amino acid sequenc
67	501.5	86.0	113	21	AA118861	Amino acid sequenc
68	501.5	86.0	113	21	AA118863	Amino acid sequenc
69	501	85.9	115	18	AA127697	Variable kappa cha
70	501	85.9	134	18	AA106206	Xenograft antibody
71	500.5	85.8	110	13	AA130012	Light chain variab
72	499.5	85.7	113	20	AA150143	Antibody F19 human
73	499.5	85.7	113	21	AA118869	Amino acid sequenc
74	499.5	85.7	240	20	AA150161	Human reshaped F19
75	499.5	85.7	432	23	AA1017498	Antibody-cytokine
76	499.5	85.7	480	23	AA1017495	Antibody-cytokine
77	499.5	85.7	601	23	AA149760	TNF-selectokine pr
78	499.5	85.7	614	23	AA1017494	Antibody-cytokine
79	499.5	85.7	658	23	AA149759	TNF-selectokine pr
80	498	85.4	112	15	AA154932	FC receptor humani
81	498	85.4	301	18	AA11507	Single chain, huma
82	498	85.4	301	20	AA173217	Multispecific sing
83	498	85.4	301	22	AA185454	Single chain human

84 498 85.4 301 22 AAB61959 Single chain human
85 498 85.4 352 20 AAY06272 Anti-Fc gamma rece
86 498 85.4 553 18 AAW11508 Single chain anti-
87 498 85.4 553 20 AAW73223 H22-anti-CEA antib
88 498 85.4 553 22 AAB85455 Bispecific single
89 498 85.4 553 22 AAB61960 Bispecific single
90 496.5 85.2 114 13 AAR22419 Murine kappa chain
91 496.5 85.2 114 15 AAR63803 Mouse HMFgl light
92 496.5 85.2 116 23 AAU81281 Human trkC antibod
93 496.5 85.2 133 21 AAY83647 Consensus contracc
94 496.5 85.2 134 14 AAR38317 Sequence encoded b
95 495.5 85.0 113 13 AAR25407 Light chain variab
96 494.5 84.8 113 21 AAB18867 Amino acid sequenc
97 494.5 84.8 244 21 AAY96305 Human IGFAM-17 imm
98 494.5 84.8 260 14 AAR38318 Sequence of single
99 490.5 84.1 220 18 AAW07528 Anti-HGF receptor
100 490 84.0 132 21 AAB07964 Amino acid sequenc

ALIGNMENTS

RESULT 1
AAR54933
ID AAR54933 standard; peptide; 112 AA.
AC AAR54933;
XX 19-OCT-1994 (first entry)
DT MAB 022 VK chain.
DE
DE
XX
KW Fc receptor; FcR; humanized antibody; hAb; IgG; cancer; allergy;
KW autoimmune disease; heteroantibody; bifunctional antibody;
KW immunotoxin; CDR; complementarity determining region; VH;
KW heavy chain variable region; VK; kappa chain variable region;
KW mononuclear phagocyte; PCR; polymerase chain reaction; primer;
KW site-directed mutagenesis; monoclonal antibody; MAb.
XX
OS Mus sp.
XX
XX WO9410332-A.
PN
XX
PD 11-MAY-1994.
XX
XX
XX 04-NOV-1993; 93WO-US10384.
PF
XX
PR 04-NOV-1992; 92GB-0023377.
PR (MEDA-) MEDAREX INC.
XX
PI Carr FJ, Harris WJ, Tempest PR;
XX
DR WPI; 1994-167486/20.
XX
XX
PT New humanised antibodies to Fc receptors - used for diagnosis or
PT for treatment of e.g. cancer, allergies and infectious and
PT auto-immune diseases
XX
PS Disclosure; Page 23; 36pp; English.
XX
CC Humanized antibodies (hAbs) for IgG Fc receptors on human phagocytes
CC comprise the CDR of mouse monoclonal antibody 22 (from hybridoma
CC 022WCL-1), VH chains from human Igs NEWM or KOL, and VK chains from
CC Ig REI. Sequences are provided for mouse 022 VH (AAR54931),
CC humanized NEWM-based VH (022 NMVH, AAR54929), humanized KOL-
CC based VH (022 KLHV, AAR54930), mouse 022 VK (AAR54933) and humanized
CC REI- based VK (022 HuVK, AAR54932). During hAb production, VH and VK
CC cDNAs were PCR amplified using primers given in AAQ65378-87.
CC Mutagenesis of clone M13VHPCR2 KOL VH (L71R) was performed using
CC oligos AAQ5388-89. The hAbs can be used in heteroantibody,
CC bifunctional antibody and immunotoxin production.
XX

SQ Sequence 112 AA;
Query Match 94.5%; Score 551; DB 15; Length 112;
Best Local Similarity 93.8%; Pred. No. 1.4e-39;
Matches 105; Conservative 5; Mismatches 2; Indels 0; Gaps 0;
QY 1 DIVMTQSPDLSAVSLGERVTMCKSSQSVLYSSNQKNYLAWYQQKPGQSPKLLIYVASTR 60
Db 1 NIVMTQSPSSLAVSAGEKVTMCKSSQSVLYSSNQKNYLAWYQQKPGQSPKLLIYMASTR 60
QY 61 ESGVPDRFSGSGGTDTFTLTISVQAEDLAVYYCHQYLSSYTFGGGTCKLEIK 112
Db 61 ESGVPDRFTGSGGTDTFTLTISVQAEDLAVYYCHQYLSSWTFGGGTCKLEIK 112
RESULT 2
ID AAW90226 standard; Protein; 263 AA.
XX
AC AAW90226;
XX
DT 10-MAY-1999 (first entry)
XX
DE Anti-B7.2 monospecific triabody 1G10.
XX
KW B7 binding molecule; costimulatory molecule; B7.1; CD80; B7.2;
KW CD86; T cell activation; inhibitor; graft versus host disease;
KW transplant rejection; allograft rejection; autoimmune disease;
KW allergy; therapy; human; triabody; antibody; 1G10.
XX
OS Chimeric - Mus sp.
OS Chimeric - synthetic.
XX
FH Key Location/Qualifiers
FT Peptide 1..24
FT /note= "pelB signal peptide"
FT Region 25..144
FT /note= "anti B7.2 MAb VH region"
FT Peptide 145..257
FT /note= "anti B7.2 MAb VL region"
FT Peptide 258..263
FT /note= "His6 tag"
XX
PN WO9858965-A2.
XX
PD 30-DEC-1998.
XX
XX 22-JUN-1998; 98WO-EP03791.
PF
XX 20-JUN-1997; 97EP-0870092.
PR
XX (INNO-) INNOGENETICS NV.
PA
XX Bosman A, Buyse M, Lorre K, Sablon E;
PI
XX WPI; 1999-105615/09.
DR N-PSDB; AAX01660.
XX
PT New molecules which bind B7.1 and B7.2 - useful to prevent and treat
PT immune diseases including allograft rejection
XX
XX Example 7.3; Fig 34; 182pp; English.
PS
CC This polypeptide comprises a 1G10 monospecific triabody composed
CC of the VH region of anti-B7.2 monoclonal antibody (MAb) 1G10
CC joined to the VL region of 1G10. A triabody is a mono- a bi- or
CC a trispecific molecule recognising simultaneously e.g. two B7.2
CC and one B7.1 molecules. It has a rigid structure that prevents
CC simultaneous binding to the 3 targets. Each antigen-binding site
CC is formed by pairing of one VH and one VL domain from the same or
CC from two different polypeptides. The invention relates to novel
CC molecules, including triabodies, which can cross-link and/or
CC cross-react with the costimulatory molecules B7.1 and B7.2 expressed

PT immune diseases including allograft rejection

XX Example 7.2; Fig 26; 182pp; English.

PS

XX This polypeptide comprises a 1G-10 monospecific diabody composed

CC of the VH region of anti-B7.2 monoclonal antibody (MAB) 1G-10

CC joined via a short, flexible linker to the VL region of 1G-10.

CC Mono- or bispecific bivalent molecules are generated by shortening

CC the flexible linker sequence between the VH and VL of the anti-B7.1

CC scFv B7-24, the anti-B7.2 scFv 1G10 and the scFv molecule with

CC dual specificity for B7.1 and B7.2 (B7.12) to 5-10 residues, and

CC for bispecific molecules by cross-pairing the VH and VL domains

CC from the 2 scFvs with different antigens to novel molecules, including

CC diabodies, which can cross-link and/or cross-react with the

CC costimulatory molecules B7.1 and B7.2 expressed on professional

CC antigen-presenting cells, leading to the inhibition of antigen-

CC specific T cell activation. Methods are provided for the

CC production of such B7-binding molecules, and for their use in the

CC treatment or prevention of diseases of the immune system, in

CC particular graft rejection, graft versus host disease, allergy and

CC autoimmune diseases (claimed).

XX

SQ Sequence 268 AA;

Query Match 92.6%; Score 540; DB 20; Length 268;

Best Local Similarity 92.0%; Pred. No. 2.9e-38;

Matches 103; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 DIVMTQSPDLSAVSLGERVTMCKSSQSVLYSSNQKNYLAAYQKPGQSPKLLIYWASTR 60

DB 150 DIELTQSPSSLAWSAGEEVTMTCKSSQSVLYSSNQKNYLAAYQKPGQSPKLLIYWASTR 209

QY 61 ESGVPDRFSGSGGTDFLTLSVQAEDLAVYYCHOYLSSYTFGGTKLEIK 112

DB 210 ESGVPDRFTGSGGTDFSLTSSVQAEDLAVYYCHOYLSSYTFGGTKLEIK 261

RESULT 5

AAW90224

ID AAW90224 standard; Protein; 273 AA.

XX

AC AAW90224;

XX

DT 10-MAY-1999 (first entry)

XX

DE Anti-B7.1/anti-B7.2 bispecific diabody II.

XX

B7 binding molecule; costimulatory molecule; B7.1; CD80; B7.2;

CD86; T cell activation; inhibitor; graft versus host disease;

transplant rejection; allograft rejection; autoimmune disease;

allergy; therapy; human; diabody; antibody; B7-24; 1G10.

XX

OS Chimeric - Mus sp.

OS Chimeric - synthetic.

XX

Key Location/Qualifiers

FT Peptide 1..39

FT /note= "g3p signal peptide"

FT 40..155

FT /note= "anti B7.1 MAB VH region"

FT 156..160

FT /note= "G4S flexible linker"

FT 161..273

FT /note= "anti B7.2 MAB VL region"

FT Misc-difference 21

FT /note= "encoded by TCA"

XX

WO9858965-A2.

XX

30-DEC-1998.

XX

22-JUN-1998; 98WO-EP03791.

PF

XX 20-JUN-1997; 97EP-0870092.

PR

XX (INNO-) INNOGENETICS NV.

PA

XX Bosman A, Buyse M, Lorre K, Sablon E;

PI

XX WPI; 1999-105615/09.

DR N-PSDB; AAX01658.

XX

PT New molecules which bind B7.1 and B7.2 - useful to prevent and treat

PT immune diseases including allograft rejection

XX

PS Example 7.2; Fig 30; 182pp; English.

XX

CC This polypeptide comprises a bispecific diabody composed of the VH

CC region of anti-B7.1 monoclonal antibody (MAB) B7-24 joined via a

CC short, flexible linker to the VL region of anti-B7.2 MAB 1G10.

CC Mono- or bispecific bivalent molecules are generated by shortening

CC the flexible linker sequence between the VH and VL of the anti-B7.1

CC scFv B7-24, the anti-B7.2 scFv 1G10 and the scFv molecule with

CC dual specificity for B7.1 and B7.2 (B7.12) to 5-10 residues, and

CC for bispecific molecules by cross-pairing the VH and VL domains

CC from the 2 scFvs with different antigen recognition (B7.1/B7.2 and

CC B7.12/B7.12). The invention relates to novel molecules, including

CC diabodies, which can cross-link and/or cross-react with the

CC costimulatory molecules B7.1 and B7.2 expressed on professional

CC antigen-presenting cells, leading to the inhibition of antigen-

CC specific T cell activation. Methods are provided for the

CC production of such B7-binding molecules, and for their use in the

CC treatment or prevention of diseases of the immune system, in

CC particular graft rejection, graft versus host disease, allergy and

CC autoimmune diseases (claimed).

XX

SQ Sequence 273 AA;

Query Match 92.6%; Score 540; DB 20; Length 273;

Best Local Similarity 92.0%; Pred. No. 2.9e-38;

Matches 103; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 DIVMTQSPDLSAVSLGERVTMCKSSQSVLYSSNQKNYLAAYQKPGQSPKLLIYWASTR 60

DB 161 DIELTQSPSSLAWSAGEEVTMTCKSSQSVLYSSNQKNYLAAYQKPGQSPKLLIYWASTR 220

QY 61 ESGVPDRFSGSGGTDFLTLSVQAEDLAVYYCHOYLSSYTFGGTKLEIK 112

DB 221 ESGVPDRFTGSGGTDFSLTSSVQAEDLAVYYCHOYLSSYTFGGTKLEIK 272

RESULT 6

AAW90218

ID AAW90218 standard; Protein; 556 AA.

XX

AC AAW90218;

XX

DT 10-MAY-1999 (first entry)

XX

DE Bispecific tetraivalent antibody BiTab1G10-B7-24H6.

XX

B7 binding molecule; costimulatory molecule; B7.1; CD80; B7.2;

CD86; T cell activation; inhibitor; graft versus host disease;

transplant rejection; allograft rejection; autoimmune disease;

allergy; therapy; human; bispecific tetraivalent antibody; BiTab;

BiTab1G10-B7-24H6.

XX

OS Chimeric - Mus sp.

OS Chimeric - Homo sapiens.

OS Chimeric - synthetic.

XX

Key Location/Qualifiers

FT Region 1..120

FT /note= "VH region anti B7.2 MAB"

FT Peptide 121..135

FT Region /note= "(G4S3) flexible linker"
FT 136..248
FT /note= "VL region anti B7.2 MAb"
FT 249..259
FT /note= "human IgG3 hinge region"
FT 260..285
FT /note= "helix-turn-helix dimerisation domain"
FT 286..305
FT /note= "human IgG3 hinge domain"
FT 306..426
FT /note= "VH region anti B7.1 MAB"
FT 427..441
FT /note= "(G4S3)flexible linker"
FT 442..550
FT /note= "VL region anti B7.1 MAB"
FT 551..556
FT /note= "His6 tag"
XX
PN WO9858965-A2.

30-DEC-1998.

22-JUN-1998; 98WO-EP03791.
XX
PR
XX 20-JUN-1997; 97EP-0870092.
XX (INNO-) INNOGENETICS NV.
XX
PI Bosman A, Buyse M, Lorre K, Sablon E;
XX
DR WPI; 1999-105615/09.
DR N-PSDB; AAX01652.
XX
PT New molecules which bind B7.1 and B7.2 - useful to prevent and treat
PT immune diseases including allograft rejection
XX
PS Example 7.1; Fig 18; 182pp; English.
XX
CC This polypeptide comprises the bispecific tetraivalent antibody
CC BiTablG10-B7-24H6. The molecule consists of 4 scFvs, i.e. 2
CC anti B7.1 scFvs and 2 anti B7.2 scFvs (tetravalency). One single
CC BiTab is a homodimer of 2 identical molecules, each containing both
CC an anti B7.1 and anti B7.2 scFv (bispecificity). An anti-B7.1 and
CC and anti-B7.2 scFv are linked using a dimerisation domain (see
CC AAW90219), which drives the homodimerisation of the molecule. DNA
CC (see AAX01652) encoding the BiTab has been constructed to allow
CC expression of the BiTab in transformed E. coli cells. The BiTab
CC cross-links, and/or cross-reacts, with the costimulatory molecules
CC B7.1 and B7.2 that are expressed on the membrane of professional
CC antigen-presenting cells, leading to the inhibition of antigen-
CC specific T cell activation. The invention relates to such
CC B7-binding molecules, methods for their production, and their use
CC for treating or preventing diseases of the immune system, in
CC particular graft rejection, graft versus host disease, allergy and
CC autoimmune diseases (claimed).
XX
SQ Sequence 556 AA:

Query Match 92.6%; Score 540; DB 20; Length 556;
Best Local Similarity 92.0%; Pred. No. 5.9e-38;
Matches 103; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Oy 1 DIVMTQSPDSLAVSLGERVTMCKSSQSVLYSSNQKNYLAWYQKPGQSPKLLIYWASTR 60
|| :||| |||| || ||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 136 DIELTQSPSSLAVSAGEEVTMTCKSSQSVLYSSNQKNYLAWYQKPGQSPKLLIYWASTR 195

Oy 61 ESGVDPREFSGSGGTDFTLTISSVQAEDLAVYYCHQYLSSYTFGGGTKLEIK 112
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 196 ESGVDPRTGSGGTDFSLTISSVQAEDLAVYYCHQYLSSYTFGGGTKLEIK 247

RESULT 7
AAW90217

ID AAW90217 standard; Protein; 580 AA.
XX
AC AAW90217;
XX
DT 10-MAY-1999 (first entry)
XX
DE Bispecific tetraivalent antibody BiTabB7-24-IG10H6.
XX
KW B7 binding molecule; costimulatory molecule; B7.1; CD80; B7.2;
KW CD86; T cell activation; inhibitor; graft versus host disease;
KW transplant rejection; allograft rejection; autoimmune disease;
KW allergy; therapy; human; bispecific tetraivalent antibody; BiTab;
KW BiTabB7-24-IG10H6.
XX
OS Chimeric - Mus sp.
OS Chimeric - Homo sapiens.
OS Chimeric - synthetic.
XX
FH Key Location/Qualifiers
FT Peptide 1..24
FT /note= "pelB signal peptide"
FT 25..138
FT /note= "VH region anti B7.1 MAB"
FT 139..153
FT /note= "(G4S3) flexible linker"
FT 154..262
FT /note= "VL region anti B7.1 MAB"
FT 263..273
FT /note= "human IgG3 hinge region"
FT 274..308
FT /note= "helix-turn-helix dimerisation domain"
FT 309..319
FT /note= "human IgG3 hinge domain"
FT 320..446
FT /note= "VH region anti B7.2 MAB"
FT 447..461
FT /note= "(G4S3)flexible linker"
FT 462..574
FT /note= "VL region anti B7.2 MAB"
FT 575..580
FT /note= "His6 tag"
FT Misc-difference 261
FT /note= "encoded by CTG"
FT Misc-difference 322..327
FT /note= "codons for these amino acids are not
FT present in the DNA sequence for
FT BiTabB7-24-IG1-H6 provided in the
FT specification"
XX
PN WO9858965-A2.
XX
PD 30-DEC-1998.
XX
PF 22-JUN-1998; 98WO-EP03791.
XX
PR 20-JUN-1997; 97EP-0870092.
XX (INNO-) INNOGENETICS NV.
XX Bosman A, Buyse M, Lorre K, Sablon E;
XX WPI; 1999-105615/09.
XX N-PSDB; AAX01651.
DR New molecules which bind B7.1 and B7.2 - useful to prevent and treat
DR immune diseases including allograft rejection
XX
PS Example 7.1; Fig 16; 182pp; English.
XX
CC This polypeptide comprises the bispecific tetraivalent antibody
CC BiTabB7-24-IG10H6. The molecule consists of 4 scFvs, i.e. 2
CC anti B7.1 scFvs and 2 anti B7.2 scFvs (tetravalency). One single
CC BiTab is a homodimer of 2 identical molecules, each containing both

```

CC an anti B7.1 and anti B7.2 scFv (bispecificity). An anti-B7.1 and
CC and anti-B7.2 scFv are linked using a dimerisation domain (see
CC AAW90219), which drives the homodimerisation of the molecule. DNA
CC (see AAX01651) encoding the BitAb has been constructed to allow
CC expression of the BitAb in transformed E. coli cells. The BitAb
CC cross-links, and/or cross-reacts, with the costimulatory molecules
CC B7.1 and B7.2 that are expressed on the membrane of professional
CC antigen-presenting cells, leading to the inhibition of antigen-
CC specific T cell activation. The invention relates to such
CC B7-binding molecules, methods for their production, and their use
CC for treating or preventing diseases of the immune system, in
CC particular graft rejection, graft versus host disease, allergy and
CC autoimmune diseases (claimed).
CC
XX
SQ Sequence 580 AA;

Query Match 92.6%; Score 540; DB 20; Length 580;
Best Local Similarity 92.0%; Pred. No. 6.1e-38;
Matches 103; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

1 DIVMTQSPDSLAVSLGERVTMCKSSQSVLYSSNOKNYLAWYQQKPGQSPKLLIYWASTR 60
|| :||| ||||| || ||| ||||| ||||| ||||| ||||| ||||| |||||
462 DIELTQSPSSLAVSAGEEVTMTCKSSQSVLYSSNOKNYLAWYQQKPGQSPKLLIYWASTR 521

Db

QY 61 ESGVPDRFGSGSGTDFTLTISVVQAEDLAVYYCHQYLSSTYTFGGGTTKLEIK 112
||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
Db 522 ESGVPDRFTGGSGTDFSLTISVVQAEDLAVYYCHQYLSSTYTFGGGTTKLEIK 573

```

XX	AC	AAR922215;			
XX	DT	28-MAY-1996	(first entry)		
XX	DE	LL2	Mab VK region.		
XX	KW	Humanised antibody;	monoclonal antibody; MAb; LL2; B-cell lymphoma;		
XX	KW	leukaemia; therapy;	diagnosis; complementarity determining region;		
XX	KW	CDR; antibody engineering.			
OS		Mus musculus.			
XX	Key	Location/Qualifiers			
FH	Region	24..40			
FT		/label= CDR1			
	Region	/note= "claim 6, page 44"			
FT		56..62			
FT		/label= CDR2			
FT	Region	/note= "claim 7, page 44"			
FT		95..103			
FT		/label= CDR3			
FT		/note= "claim 8, page 44"			
XX					
PN		WO9604925-A1.			
XX					
PD		22-FEB-1996.			
XX					
PF		11-AUG-1995;	95WO-US09641.		
XX					
PR		12-AUG-1994;	94US-0289576.		
XX					
PA		(IMMU-)	IMMUNOMEDICS INC.		
XX					
PI		Hansen H,	Leung S;		
XX					
DR		WPI;	1996-139454/14.		
DR		N-PSDB;	AAT15802.		
XX					
PT		Chimeric and humanised LL2 antibodies -	used to produce conjugates		
PT		for the therapy and diagnosis of B-cell lymphoma(s) and			

```

leukaemia(s).
Claim 5; Page 35-36; 70pp; English.

The complementarity determining regions (CDRs) of mouse monoclonal antibody (MAB) LL2 VK (AAR92215) and VH (AAR92216) regions were recombinantly linked to the framework sequences of human VK and VH regions, respectively, to give humanised LL2 VK (AAR92217) and VH (AAR92218). These were subsequently linked, respectively, to human kappa and IgG1 constant regions. A humanised MAB was obtd. that retained the B-lymphoma and leukaemia cell targeting and internalisation characteristics of the parental LL2 MAb, and which exhibited a lowered HAMA reaction. It can be linked to e.g. a cytostatic agent for therapeutic appln.

Sequence 113 AA;

Query Match 90.4%; Score 527; DB 17; Length 113;
Best Local Similarity 88.4%; Pred. No. 1.6e-37;
Matches 99; Conservative 6; Mismatches 7; Indels 0; Gaps 0;

QY 1 DIVMTQSPDLSAVSLGERVTMCKSSQSVLYSSNQKNYLAWYQQKPGQSPKLLIYWASTR 60
Db 1 DIQLTQSPSSLAWSAGENVMTSCKSSQSVLYSANHKNYLAWYQQKPGQSPKLLIYWASTR 60
QY 61 ESGVPDRFRSGSGGTDFTLTISVVQAE DLAVYYCHQYLSSTYFGGGTKLEIK 112
Db 61 ESGVPDRFTGSGSGTDFTLTISRQVEDLAIYYCHQYLSSTYFGGGTKLEIK 112

RESULT 9
AAW27695
ID AAW27695 standard; Protein; 113 AA.
XX
AC AAW27695;
XX
DT 14-APR-1998 (first entry)
XX
DE Variable kappa chain of MAB LL2.
XX
KW Variable kappa chain; B cell; monoclonal antibody; MAB; LL2;
KW B cell lymphoma; lymphocytic leukaemia cell; murine;
KW diagnosis; treatment; B cell malignancy; non-Hodgkins lymphoma;
KW chronic lymphocytic leukaemia.
XX
OS Mus sp.
XX Synthetic.
XX
FH Location/Qualifiers
FT Region 18..20
FT /note= "potential N-linked glycosylation site"
FT Region 24..40
FT /note= "complementarity determining region 1"
FT Region 56..62
FT /note= "complementarity determining region 2"
FT Region 95..102
FT /note= "complementarity determining region 3"
XX
XX WO9734632-A1.
XX
XX 25-SEP-1997.
PD
XX
XX 19-MAR-1997; 97WO-US04196.
PF
XX 20-MAR-1996; 96US-0013709.
PR
XX (IMMU-) IMMUNOMEDICS INC.
XX
XX Hansen H, Leung S, Qu Z;
XX
XX WPI; 1997-479995/44.
XX
XX N-PSDB; AAT88128.
XX

```

PT Monoclonal antibody engineered to contain glycosylation site - in
PT non-Fc constant heavy or light chain region, useful to diagnose or
PT treat B cell malignancies, e.g. non-Hodgkins lymphoma
XX
PS Example 3; Fig 4A; 88pp; English.
XX
CC The present sequence is the variable kappa chain of the
CC B cell specific monoclonal antibody (Mab) LL2, which contains an
CC engineered tri-peptide N-glycan acceptor sequence. LL2 is a highly
CC specific anti-B cell lymphoma and anti-lymphocytic leukaemia cell
CC murine Mab. The Mab can be used to diagnose or treat B
CC cell malignancies, e.g. non-Hodgkins lymphoma or chronic
CC lymphocytic leukaemia. The glycosylation site allows a label or
CC therapeutic agent of increased size to be conjugated to the
CC carbohydrate moiety, without affecting the Mab's binding affinity
CC or specificity.
XX
SQ Sequence 113 AA;
Query Match 90.4%; Score 527; DB 18; Length 113;
Best Local Similarity 88.4%; Pred. No. 1.6e-37;
Matches 99; Conservative 6; Mismatches 7; Indels 0; Gaps 0;
QY 1 DIVMTQSPDSLAVSLGERVTMCKSSQSVLYSSNQKNYLAHYQKPGQSPKLLIYWASTR 60
DB 1 DIQLTQSPSSLAVSAGENVMTMCKSSQSVLYSSNQKNYLAHYQKPGQSPKLLIYWASTR 60
QY 61 ESGVPDRFSGSGGTFTLTISVQAEDLAVYYCHQYLSSYTFGGGKLEIK 112
DB 61 ESGVPDRFSGSGGTFTLTISRQVEDLAIYYCHQYLSSYTFGGGKLEIK 112
RESULT 10
AAB62769
ID AAB62769 standard; Protein; 113 AA.
XX
AC AAB62769;
XX
DT 03-APR-2001 (first entry)
XX
DE Human HIV-1 monoclonal antibody SEQ ID NO: 68.
XX
KW Human immunodeficiency virus-1; HIV-1; human monoclonal antibody;
KW envelope glycoprotein; gp120; diagnosis.
XX
OS Homo sapiens.
XX
WO200100678-A1.
04-JAN-2001.
23-JUN-2000; 2000WO-US17327.
30-JUN-1999; 99US-0141701.
(USSH) US DEPT HEALTH & HUMAN SERVICES.
Watkins BA, Reitz MS;
WPI; 2001-112438/12.
N-PSDB; AAF29070.
Novel human monoclonal antibody immunoreactive with human
immunodeficiency virus-1 glycoprotein gp120, useful for detecting HIV-1
in biological sample and providing passive immunotherapy to HIV-1
infected mammal -
Claim 1; Page 65; 81pp; English.
The present invention provides the protein and coding sequences for the
variable regions of human monoclonal antibodies which are immunoreactive
with human immunodeficiency virus-1 (HIV-1) envelope glycoprotein gp120.
These can be used in diagnosis and therapy of HIV-1 infection.

XX SQ Sequence 113 AA;
Query Match 90.2%; Score 526; DB 22; Length 113;
Best Local Similarity 88.4%; Pred. No. 1.9e-37;
Matches 99; Conservative 7; Mismatches 6; Indels 0; Gaps 0;
QY 1 DIVMTQSPDSLAVSLGERVTMCKSSQSVLYSSNQKNYLAHYQKPGQSPKLLIYWASTR 60
DB 1 DLVMTQSPDSLAVSLGERATINCKSSQSVLYSSNNKNYLAHYQKPGQPPKLLIYWASTR 60
QY 61 ESGVPDRFSGSGGTFTLTISVQAEDLAVYYCHQYLSSYTFGGGKLEIK 112
DB 61 ESGVPDRFSGSGGTFTLTISLQAEDVAVYYCQYFNTPTFFGGGKVEIK 112
RESULT 11
AAR50313
ID AAR50313 standard; Protein; 113 AA.
XX
AC AAR50313;
XX
DT 05-OCT-1994 (first entry)
XX
DE Humanised light chain variable region Pfhlzcl-1.
XX
KW Monoclonal antibody; Plasmodium falciparum; CDR;
KW complementarity determining region; fusion protein;
KW murine; variable; light; heavy; chain; malaria.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Region 24..40
FT /label= CDR1
FT Region 56..62
FT /label= CDR2
FT Region 95..103
FT /label= CDR3
XX
PN WO9405690-A.
XX
PD 17-MAR-1994.
XX
PF 08-SEP-1993; 93WO-US08435.
XX
PR 09-SEP-1992; 92US-0941654.
XX
PA (SMIK) SMITHKLINE BEECHAM CORP.
PA (USSA) US SEC OF ARMY.
PA (USNA) US SEC OF NAVY.
XX
PI Charoenvit Y, Hoffman S, Hurle M, Rosenberg M;
PI Sadoff JC, Sylvester DR, Gross MS;
XX
WPI; 1994-101115/12.
DR N-PSDB; AAQ44827.
XX
PT New engineered antibodies and fusion proteins for preventing
PT Plasmodium infection - contg. murine antibody CDR sequences, and
PT corresp. nucleic acid, vectors and transformed cells
XX
PS Claim 5; Fig 2; 98pp; English.
XX
CC Naturally-occurring (AAQ44841-42) and synthetic (AAQ44825-28)
CC variable light chain and variable heavy chain sequences
CC derived from murine mAb NFS2 are provided. Murine mAb NFS2,
CC its variable chain peptides, CDRs, functional fragments,
CC Fab fragments, and analogs are useful in prodn. of fusion
CC proteins, esp. engineered antibodies. These prods. are
CC used to protect humans against Plasmodium infections.
XX
SQ Sequence 113 AA;

Query Match 90.1%; Score 525.5; DB 15; Length 113;
Best Local Similarity 90.3%; Pred. No. 2.1e-37;
Matches 102; Conservative 5; Mismatches 5; Indels 1; Gaps 1;

QY 1 DIVMTQSPDSLAVSLGERVTMCKSSQSVLYSSNQKNYLAHYQQKPGQSPKLLIYWASTR 60
Db 1 DIVMTQSPDSLAVSLGERATINCKSSQSVLYSSNQKNYLAHYQQKPGQSPKLLIYWASTR 60

QY 61 ESGVPDRFSGSGGTDTLTISVQAEDLAVYCHQYLS-SYTFGGGKLEIK 112
Db 61 ESGVPDRFSGSGGTDTLTISVQAEDLAVYCHQYLSYPRTEFGGKVEIK 113

RESULT 12
AAG65565
ID AAG65565 standard; protein; 120 AA;
XX
AC AAG65565;
XX
30-NOV-2001 (first entry)
DE Amino acid sequence of protein seq Id No. 90.
XX
KW Gene library; immunoglobulin; antibody library; human.
XX
OS Homo sapiens.
XX
PN WO200162907-A1.
XX
PD 30-AUG-2001.
XX
PF 22-FEB-2001; 2001WO-JP01298.
XX
PR 22-FEB-2000; 2000JP-0050543.
XX
PA (MEDI-) MEDICAL & BIOLOGICAL LAB CO LTD.
XX
PI Kurosawa Y, Akahori Y, Iba Y, Morino K, Shinohara M, Takahashi M;
PI Okuno Y, Shiraki K;
XX
DR WPI; 2001-565420/63.
DR N-PSDB; AAH47729.
XX
PT Producing gene libraries and antibody libraries, involves selecting a
PT light chain that binds to a heavy chain product to produce a functional
PT formation, and producing a gene library of the light chain variable
PT regions -
XX
SQ Examples; p 165; 181pp; Japanese.

CC The invention relates to producing gene libraries, comprising
CC immunoglobulin light and heavy variable region. The method involves
CC selecting light chain that binds with the heavy chain product to produce
CC a functional conformation, producing a gene library comprising a
CC collection of these light chain variable genes, and combining with gene
CC library of heavy chain variable genes. The method is used for production
CC of gene and antibody libraries.

QY 1 DIVMTQSPDSLAVSLGERVTMCKSSQSVLYSSNQKNYLAHYQQKPGQSPKLLIYWASTR 60
Db 1 DIVMTQSPDSLAVSLGERATINCKSSQSVLYSSNQKNYLAHYQQKPGQSPKLLIYWASTR 60

QY 61 ESGVPDRFSGSGGTDTLTISVQAEDLAVYCHQYLS-SYTFGGGKLEIK 112
Db 61 ESGVPDRFSGSGGTDTLTISVQAEDLAVYCHQYLSYPRTEFGGKVEIK 113

RESULT 13
AAR28807
ID AAR28807 standard; Protein; 112 AA.
XX
AC AAR28807;
XX
DT 02-APR-1993 (first entry)
XX
DE 5A8 VL.
XX
KW Primer; polymerase chain reaction; amplify; 5A8; heavy chain;
KW variable; antibody; VH1BACK; VH1FOR; VH01; light chain; humanised;
KW gp120; CD4; cell surface glycoprotein; CD4+; lymphocyte;
KW HIV-induced syncytia formation.
XX
OS Synthetic.
XX
PN WO9209305-A.
XX
PD 11-JUN-1992.
XX
PF 27-NOV-1991; 91WO-US08843.
XX
PR 27-NOV-1990; 90US-0618542.
XX
PA (BIOJ) BIOGEN INC.
XX
PI Burkly LC, Chisholm PL, Rosa JJ, Rosa MD, Thomas DW;
XX
DR WPI; 1992-398399/48.
DR N-PSDB; AAQ30881.
XX
PT New anti-CD4 antibody homologues - which bind CD4, do not block
PT binding of HIV gp120 to CD4 but block HIV-induced syncytia
PT formation between CD4+ cells
XX
PS Claim 27; Page 129-30; 205pp; English.

CC The sequence represents the 5A8 light chain variable region (5A8 VL)
CC of the antibody of the invention. This sequence was inserted into
CC the vector pUC19 aligned with the pMDR927 insert. The encoding
CC DNA was amplified using the primer sequences given in AAQ30569-70. 5A8
CC VL was used in conjunction with the 5A8 heavy chain variable region and
CC these two chains were then humanised. The antibody produced was shown
CC to bind to CD4 but did not block the binding of gp120 to CD4. CD4 is a
CC cell surface glycoprotein of CD4+ lymphocytes (helper/inducer cells).
CC The antibody blocked HIV-induced syncytia formation. This antibody
CC can be used in the detection, prophylaxis and treatment of diseases
CC caused by infective agents whose primary targets are CD4+ cells.

XX
SQ Sequence 112 AA;

Query Match 89.4%; Score 521; DB 13; Length 112;
Best Local Similarity 89.3%; Pred. No. 5e-37;
Matches 100; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

QY 1 DIVMTQSPDSLAVSLGERVTMCKSSQSVLYSSNQKNYLAHYQQKPGQSPKLLIYWASTR 60
Db 1 DIVMTQSPDSLAVSLGERVTMCKSSQSVLYSSNQKNYLAHYQQKPGQSPKLLIYWASTR 60

QY 61 ESGVPDRFSGSGGTDTLTISVQAEDLAVYCHQYLS-SYTFGGGKLEIK 112
Db 61 ESGVPDRFSGSGGTDTLTISVQAEDLAVYCHQYLSYPRTEFGGKVEIK 113

RESULT 14
AAR50314
ID AAR50314 standard; Protein; 113 AA.
XX
AC AAR50314;
XX
DT 05-OCT-1994 (first entry)

XX DE Humanised light chain variable region Pfhlz1c1'-2.
XX KW Monoclonal antibody; Plasmodium falciparum; CDR;
KW complementarity determining region; fusion protein;
KW murine; variable; light; heavy; chain; malaria.
XX OS Synthetic.

XX FH Key Location/Qualifiers
XX FT Region 24..40
FT /label= CDR1
FT Region. 56..62
FT /label= CDR2
FT Region 95..103
FT /label= CDR3

XX PN WO9405690-A.

XX 17-MAR-1994.

XX PF 08-SEP-1993; 93WO-US08435.

XX PR 09-SEP-1992; 92US-0941654.

XX PA (SMIK) SMITHKLINE BEECHAM CORP.
PA (USSA) US SEC OF ARMY.
PA (USNA) US SEC OF NAVY.

XX PI Charoenvit Y, Hoffman S, Hurle M, Rosenberg M;
PI Sadoff JC, Sylvester DR, Gross MS;
XX N-PSDB; AAQ44828.

XX DR WPI; 1994-101115/12.

XX DR N-PSDB; AAQ44828.

XX PT New engineered antibodies and fusion proteins for preventing
PT Plasmodium infection - contg. murine antibody CDR sequences, and
PT corresp. nucleic acid, vectors and transformed cells

XX PS Claim 5; Fig 3; 98pp; English.

XX CC Naturally-occurring (AAQ44841-42) and synthetic (AAQ44825-28)
CC variable light chain and variable heavy chain sequences
CC derived from murine mAb NFS2 are provided. Murine mAb NFS2,
CC its variable chain peptides, CDRs, functional fragments,
CC Fab fragments, and analogs are useful in prodn. of fusion
CC proteins, esp. engineered antibodies. These prods. are
CC used to protect humans against Plasmodium infections.

XX Sequence 113 AA;

Query Match 89.3%; Score 520.5; DB 15; Length 113;
Best Local Similarity 89.4%; Pred. No. 5.5e-37;
Matches 101; Conservative 5; Mismatches 6; Indels 1; Gaps 1;

QY 1 DIVMTQSPDSLAVSLGERVTMCKSSQSVLYSSNQKNYLAQQKPGQSPKLLIYWASTR 60

Db 1 DIVMTQSPDSLAVSLGERATINCKSSQSVLYSSNQKNYLAQQKPGQSPKLLIYWASTR 60

QY 61 ESGVPDRFSGSGGTDFTLTISSVQAEDLAVYYCHQYLS-SYTFGGGTKLEIK 112

Db 61 ESGVPDRFSGSGGTDFTLTISSVQAEDVAVYYCQYYSYPRTFGGGTKVEIK 113

RESULT 15

AAW32483

ID AAW32483 standard; Protein; 155 AA.

XX AAW32483;

AC 27-MAR-1998 (first entry)

XX Kappa light chain variable region.

XX KW Catalytic antibody; growth factor; B-cell mitogenesis; LHL.seq;
KW L protein; hen egg lysozyme; HEL; kappa light chain; LEN.

XX OS Chimeric - Homo sapiens.
OS Chimeric - synthetic.

XX FH Key Location/Qualifiers
XX FT Peptide 22..29
FT /note= "FLAG epitope"
FT Peptide 147..155
FT /note= "strept-tag"

XX PN WO9735887-A1.

XX PD 02-OCT-1997.

XX PF 26-MAR-1997; 97WO-AU00194.

XX PR 27-FEB-1997; 97AU-0005375.

XX PR 26-MAR-1996; 96AU-0008951.

XX PA (AMRA-) AMRAD OPERATIONS PTY LTD.

XX PI Koentgen E, Suess GM, Tarlinton DM, Treutlein HR;

XX WPI; 1997-489572/45.

XX N-PSDB; AAW32483.

XX PT New catalytic antibody precursors - comprising a B-cell surface
PT molecule binding portion which can induce B-cell mitogenesis
XX Example 7; Page 73-74; 109pp; English.

XX CC This polypeptide comprises the variable kappa light chain from
CC human myeloma protein LEN, flanked by an N-terminal FLAG epitope
CC (see AAW32484) and a C-terminal strep-tag (see AAW32485) to facilitate
CC purification from E. coli host cells. Kappa peptide is soluble at
CC relatively high concentrations and binds to protein L. Growth
CC factor TLHL (see AAW32482) was generated from LHL (see AAW32479), kappa
CC and linker sequences. It was designed so that the kappa portion of
CC the protein could be cleaved off by TEV protease, generating LHL.
CC Production of catalytic antibodies to a specific antigen comprises
CC administering to an animal a growth factor comprising an antigen
CC capable of interacting with a B cell bound catalytic antibody. The
CC antigen is fused to a B cell surface molecule binding protein for
CC the antigen to be cleaved and for the remainder of the molecule to
CC induce B cell mitogenesis (claimed). The catalytic antibodies
CC generated by the process can have diagnostic and therapeutic
CC applications.

XX SQ Sequence 155 AA;

Query Match 89.3%; Score 520.5; DB 18; Length 155;
Best Local Similarity 89.4%; Pred. No. 7.5e-37;
Matches 101; Conservative 5; Mismatches 6; Indels 1; Gaps 1;

QY 1 DIVMTQSPDSLAVSLGERVTMCKSSQSVLYSSNQKNYLAQQKPGQSPKLLIYWASTR 60

Db 30 DIVMTQSPDSLAVSLGERATINCKSSQSVLYSSNQKNYLAQQKPGQSPKLLIYWASTR 89

QY 61 ESGVPDRFSGSGGTDFTLTISSVQAEDLAVYYCHQYLS-YTFGGGTKLEIK 112

Db 90 ESGVPDRFSGSGGTDFTLTISSVQAEDVAVYYCQYYSYTPSYFSGGTGKLEIK 142

RESULT 16

AAW06912

ID AAY06912 standard; Protein; 155 AA.

XX AAY06912;

DT 01-JUL-1999 (first entry)

XX DE Human variable kappa light chain protein.

XX KW Growth factor precursor; B-cell surface; T cell surface; CA; hepatitis;

KW catalytic antibody; immunoglobulin; tumour necrosis factor; influenza;

KW rheumatoid arthritis; toxic shock syndrome; viral docking receptor;

KW human immune virus; tumour-specific antigen; amyloid plaque; myeloma;

KW Alzheimer's disease; IGE; allergy; asthma; drug detoxification; LEN;

KW autoimmune; inflammatory disease; gene therapy; human; myeloma protein;

XX kappa light chain protein.

OS Homo sapiens.

XX WO9915563-A1.

PN 01-APR-1999.

PD 18-SEP-1998; 98WO-AU00783.

XX 19-SEP-1997; 97AU-0009306.

PR (AMRA-) AMRAD OPERATIONS PTY LTD.

PI Koentgen F, Suess GM, Tarlinton DM, Treutlein HR;

XX WPI; 1999-244394/20.

DR N-PSDB; AAX34594.

XX Growth factor precursor cleaved by antigen-specific catalytic

PT antibody

XX Example 7; Page 72-73; 101pp; English.

PS The invention relates to a growth factor precursor that comprises B-cell

XX surface binding part, T cell surface binding part, antigen cleavable

CC by a catalytic antibody (CA); and a peptide comprising heavy and light

CC chains of immunoglobulin. When the antigen is cleaved the B cell surface

CC part can interact with its target. The growth factor precursors are used

CC to select B cells that produce Ag-specific CA, and to generate CA from

CC such cells (by inducing mitogenesis, caused by the growth factor released

CC by specific cleavage). The Ag-specific CA can be directed against, e.g.

CC tumour necrosis factor (treatment of rheumatoid arthritis or toxic shock

CC syndrome); viral docking receptors (treatment of human immune virus,

CC hepatitis and influenza infections); tumour-specific antigens; amyloid

CC plaque (treatment of Alzheimer's disease or myeloma) or IGE (treatment of

CC allergies such as asthma). CA may also be used for drug detoxification,

CC to treat autoimmune or inflammatory diseases and to eliminate

CC environmental or industrial pollutants, such as plastics and petroleum.

CC Particularly the growth factor precursors are produced by delivering the

CC corresponding nucleic acid in a viral or other gene therapy vector. The

CC present sequence represents the amino acid sequence of the human variable

CC kappa light chain. This was generated based on the human myeloma protein

CC LEN sequence.

XX Sequence 155 AA;

SQ Query Match 89.3%; Score 520.5; DB 20; Length 155;

Best Local Similarity 89.4%; Pred. No. 7.5e-37;

Matches 101; Conservative 5; Mismatches 6; Indels 1; Gaps 1;

QY 1 DIVMTQSPDSLAVSLGERVTMCKSSQSVLYSSNKNYLAAYQKPGQSPKLLIYWASTR 60

Db 30 DIVMTQSPDSLAVSLGERVTMCKSSQSVLYSSNKNYLAAYQKPGQSPKLLIYWASTR 89

QY 61 ESGVPDRFSGSGGTDFLTITSSVQAEDLAVYYCHQYLSS-YTFGGGKLEIK 112

Db 90 ESGVPDRFSGSGGTDFLTITSSVQAEDLAVYYCHQYLSS-YTFGGGKLEIK 142

RESULT 17

AAW32482

ID AAW32482 standard; Protein; 342 AA.

XX

AC AAW32482;

XX 27-MAR-1998 (first entry)

XX Growth factor TLHL (catalytic antibody precursor).

DE Catalytic antibody; growth factor; B-cell mitogenesis; TLHL;

XX L protein; hen egg lysozyme; kappa light chain; ss.

KW Chimeric - Peptostreptococcus magnus.

KW Chimeric - Gallus sp.

OS Chimeric - Homo sapiens.

XX WO9735887-A1.

PN 02-OCT-1997.

XX 26-MAR-1997; 97WO-AU00194.

PF 27-FEB-1997; 97AU-0005375.

XX 26-MAR-1996; 96AU-0008951.

PR (AMRA-) AMRAD OPERATIONS PTY LTD.

XX Koentgen F, Suess GM, Tarlinton DM, Treutlein HR;

PI WPI; 1997-489572/45.

XX N-PSDB; AAT91589.

DR New catalytic antibody precursors - comprising a B-cell surface

DR molecule binding portion which can induce B-cell mitogenesis

XX Example 11; Page 66-68; 109pp; English.

PS This protein comprises growth factor TLHL, where L is the

XX immunoglobulin binding entity from Peptostreptococcus magnus, H is

CC residues 42-62 of hen egg lysozyme, and T represents the variable

CC kappa light chain (see AAW32483) from human myeloma protein LEN. It

CC was expressed in E. coli DH10B cells utilising a DNA construct (see

CC AAT91589) produced from LHL (see AAT91986), kappa (see AAT91590) and

CC synthetic linker oligonucleotides, and was purified from total

CC bacterial lysate via a strep-tag. The linker contains a cleavage

CC site for tobacco etch virus (TEV) protease. TLHL was designed so

CC that the kappa portion of the protein is cleaved by TEV protease

CC into 'T' and 'LHL' moieties. Production of catalytic antibodies to

CC a specific antigen comprises administering to an animal a growth

CC factor comprising an antigen capable of interacting with a B cell

CC bound catalytic antibody. The antigen is fused to a B cell surface

CC molecule binding protein for the antigen to be cleaved and for the

CC remainder of the molecule to induce B cell mitogenesis (claimed).

CC LHL crosslinks with surface immunoglobulin on B cells. This induces

CC B cell activation and blast formation. Internalisation and

CC processing of LHL leads to the presentation of the H peptide on

CC MHC II. T cell recognition of MHC II with the H peptide signals the

CC activated B cell to proliferate and undergo antibody class switching

CC and secretion. The catalytic antibodies generated by the process

CC can have diagnostic and therapeutic applications.

XX Sequence 342 AA;

SQ Query Match 89.3%; Score 520.5; DB 18; Length 342;

Best Local Similarity 89.4%; Pred. No. 1.6e-36;

Matches 101; Conservative 5; Mismatches 6; Indels 1; Gaps 1;

QY 1 DIVMTQSPDSLAVSLGERVTMCKSSQSVLYSSNKNYLAAYQKPGQSPKLLIYWASTR 60

Db 30 DIVMTQSPDSLAVSLGERVTMCKSSQSVLYSSNKNYLAAYQKPGQSPKLLIYWASTR 89

QY 61 ESGVPDRFSGSGGTDFLTITSSVQAEDLAVYYCHQYLSS-YTFGGGKLEIK 112

Db 90 ESGVPDRFSGSGGTDFLTITSSVQAEDLAVYYCHQYLSS-YTFGGGKLEIK 142

[illegible]

QY 1 DIVMTQSPDLSAVSLGERVTMCKSSQSVLYSSNQKNYLAWYQQKPGQSPKLLIYWASTR 60
Db 30 DIVMTQSPDLSAVSLGERATINCKSSQSVLYSSNSKNYLAWYQQKPGQPPKLLIYWASTR 89
QY 61 ESGVPDRFSGSGGTDFTLTISSVQAEDLAVYYCHQYLSS-YTFGGGTKLEIK 112
Db 90 ESGVPDRFSGSGGTDFTLTISSLQAEDVAVYYCQYYSTPYSPYSGQGTKEIK 142

RESULT 20
AAY06908
ID AAY06908 standard; Protein; 495 AA.

XX AAY06908;

XX 01-JUL-1999 (first entry)

XX CATAB-TEV aminoacid sequence.

OS Growth factor precursor; B-cell surface; T cell surface; CAB; hepatitis;
OS catalytic antibody; immunoglobulin; tumour necrosis factor; influenza;
XX rheumatoid arthritis; toxic shock syndrome; viral docking receptor;
XX human immune virus; tumour-specific antigen; amyloid plaque; myeloma;
KW Alzheimer's disease; IGE; allergy; asthma; drug detoxification;
KW autoimmune; inflammatory disease; gene therapy; protein L; P. magnus;
KW hen egg lysosome; HEL; LHL; CATAB-TEV; tobacco etch virus; TLHL.

XX Peptostreptococcus magnus.
OS Gallus sp.

XX W09915563-A1.

PD 01-APR-1999.

XX 18-SEP-1998; 98WO-AU00783.

PR 19-SEP-1997; 97AU-0009306.

XX (AMRA-) AMRAD OPERATIONS PTY LTD.

PI Koentgen F, Suess GM, Tarlinton DM, Treutlein HR;

XX WPI; 1999-244394/20.

DR N-PSDB; AAX34591.

XX Growth factor precursor cleaved by antigen-specific catalytic
PT antibody

XX Example 15; Page 64-66; 101pp; English.

CC The invention relates to a growth factor precursor that comprises B-cell
CC surface binding part, T cell surface binding part, antigen cleavable
CC by a catalytic antibody (CAB); and a peptide comprising heavy and light
CC chains of immunoglobulin. When the antigen is cleaved the B cell surface
CC part can interact with its target. The growth factor precursors are used
CC to select B cells that produce Ag-specific CAB, and to generate CAB from
CC such cells (by inducing mitogenesis, caused by the growth factor released
CC by specific cleavage). The Ag-specific CAB can be directed against, e.g.
CC tumour necrosis factor (treatment of rheumatoid arthritis or toxic shock
CC syndrome); viral docking receptors (treatment of human immune virus,
CC hepatitis and influenza infections); tumour-specific antigens; amyloid
CC plaque (treatment of Alzheimer's disease or myeloma) or IGE (treatment of
CC allergies such as asthma). CAB may also be used for drug detoxification,
CC to treat autoimmune or inflammatory diseases and to eliminate
CC environmental or industrial pollutants, such as plastics and petroleum.
CC Particularly the growth factor precursors are produced by delivering the
CC corresponding nucleic acid in a viral or other gene therapy vector. The
CC present sequence represents the amino acid sequence of CATAB-TEV which
CC is assembled from TLHL and kappa by PCR. The LHL is a growth factor
CC comprising H flanked by two L molecules where L is a B cell surface
CC molecule binding portion from protein L of P. magnus; H is a T cell
CC surface molecule binding portion from hen egg lysosome (HEL). TLHL is

CC LHL linked to kappa chain via TEV (tobacco etch virus) sensitive
CC peptide and g attached to N terminus region.

XX Sequence 495 AA;

QY Query Match 89.3%; Score 520.5; DB 20; Length 495;
Best Local Similarity 89.4%; Pred. No. 2.4e-36;
Matches 101; Conservative 5; Mismatches 6; Indels 1; Gaps 1;

QY 1 DIVMTQSPDLSAVSLGERVTMCKSSQSVLYSSNQKNYLAWYQQKPGQSPKLLIYWASTR 60
Db 30 DIVMTQSPDLSAVSLGERATINCKSSQSVLYSSNSKNYLAWYQQKPGQPPKLLIYWASTR 89
QY 61 ESGVPDRFSGSGGTDFTLTISSVQAEDLAVYYCHQYLSS-YTFGGGTKLEIK 112
Db 90 ESGVPDRFSGSGGTDFTLTISSLQAEDVAVYYCQYYSTPYSPYSGQGTKEIK 142

RESULT 21
AAY95243

ID AAY95243 standard; Protein; 137 AA.

XX AAY95243;

XX 29-AUG-2000 (first entry)

XX Humanised antibody HuCC49 light chain variable region.

KW Humanised antibody; monoclonal antibody; CC49; HuCC49; CDR;
KW complementarity determining region; mouse; human; carcinoma;
KW colon cancer; tumor associated glycoprotein-72; TAG-72;
KW tumour marker; diagnosis; therapy.

XX Chimeric - Mus musculus.

OS Chimeric - Homo sapiens.

XX Key Location/Qualifiers
FH Region 44..59
FT /note= "CDR1"
FT Region 76..82
FT /note= "CDR2"
FT Region 115..123
FT /note= "CDR3"

XX W0200026394-A1.

XX 11-MAY-2000.

XX 29-OCT-1999; 99WO-US25552.

PR 31-OCT-1998; 98US-0106534.

PR 02-NOV-1998; 98US-0106757.

XX (USSH) US DEPT HEALTH & HUMAN SERVICES.

PI Kashmiri SVS, Padlan EA, Schlom J;

XX WPI; 2000-365637/31.

XX Chimeric variants of CC49 monoclonal antibodies useful for detecting
PT and treating cancers associated with the expression of the pancreatic
PT tumor-associated antigen TAG-72 -

PS Disclosure; Fig 4; 76pp; English.

XX The present sequence is that of the light chain variable region
CC (VL) of huCC49, a humanised monoclonal antibody (MAB) formed by
CC grafting hypervariable regions from murine MAB CC49 into VL and VH
CC frameworks of human MABs LEN and 21/28' CL, respectively, while
CC retaining murine framework residues required for integrity of the
CC antigen combining site structure. HuCC49 binds to the human
CC pancreaticoma tumor associated glycoprotein-72 (TAG-72), which is
CC found on the surface of certain human tumours. The invention is

PI	Spiegel H, Zimmerman S, Emans N, Holzem A;	PI	01-DEC-1987; 87EP-0117760.
XX	WPI; 2000-339692/29.	XX	03-DEC-1986; 86JP-0288340.
DR		PR	26-NOV-1987; 87JP-0298513.
XX	New fusion proteins and gene constructs for expressing agents	XX	(SUMO) SUMITOMO CHEM IND KK.
PT	(antibodies, enzymes, vectors or molecular pathogenicides), useful for	PA	
PT	protecting plants against pathogens and increasing resistance to	XX	
PT	disease -	PI	Nakatani T, Nomura N, Horigome K, Noguchi H;
XX		XX	WPI; 1988-156310/23.
PS	Example 6; Flg 23c; 193pp; English.	DR	N-PSDB; AAN80499.
XX		XX	New gene encoding for antibody to Pseudomonas aeruginosa exotoxin -
CC	The present invention describes a fusion protein (I) comprising at least	PT	plus recombinant vectors and host cells, useful for treating
CC	one binding domain specifically recognising an epitope of a plant	PT	infections.
CC	pathogen and at least one further domain comprising a protein or peptide	XX	Claim 4; Page 25; 39pp; English.
CC	sequence which is toxic to the pathogen or detrimental to its	PS	
CC	replication, transmission or life cycle. Also described is a	CC	Sequence is variable region of light chain of anti-exotoxin
CC	pathogenicide (II) comprising (I) and a cellular targeting sequence	CC	antibody with signal sequence. AAN80498 encodes the same sequence
CC	and/or membrane localisation sequence and/or motif that leads to	CC	except that its signal peptide-encoding sequence contains an
CC	membrane anchoring; or at least one binding domain that specifically	CC	intron.
CC	recognises a viral movement and/or replicase protein. The fusion	CC	See also AAN80495-N80496, AAN80498 and AAN80941-2.
CC	protein, pathogenicide, polynucleotide, vectors, and compositions from	XX	
CC	the present invention are useful for the protection of a plant against	SQ	Sequence 133 AA;
CC	the action of a pathogen. The kit from the present invention is useful		Query Match 87.9%; Score 512.5; DB 9; Length 133;
CC	for carrying out the methods and may be employed in different		Best Local Similarity 88.5%; Pred. No. 3.1e-36;
CC	applications, for example in the diagnostic field or as research tools.		Matches 100; Conservative 5; Mismatches 7; Indels 1; Gaps 1;
CC	The kit or its components, such as the fusion protein, pathogenicide,		
CC	polynucleotides, vectors or compositions are useful in plant cell and	QY	1 DIVMTQSPDLSAVSLGERVTMCKSSQSVLYSSNQKNYLAWYQKPGQSPKLLIYWASTR 60
CC	plant tissue culture, in agriculture. They are extremely useful for		
CC	breeding new varieties of plants that display improved properties such as	Db	21 DIVMTQSPDLSAVSLGERATINCKSSQSVLYSSNNKNYLAWYQKPGQPPKLLIYWASTR 80
CC	resistance to pathogens. AAA56587 to AAA56702 and AAB09774 to B097820		
CC	represent sequences used in the exemplification of the present	QY	61 ESGVPDRFSGSGGTDTFTLTISVQAEDLAVYYCHQYLSS-YTFGGGTTKLEIK 112
CC	invention.		
XX		Db	81 ESGVPDRFSGSGGTDTFTLTISLQAEDVAVYYCQYYSTPRTFGGGTTKVEIK 133
SQ	Sequence 259 AA;		
	Query Match 88.0%; Score 513; DB 21; Length 259;	RESULT 25	
	Best Local Similarity 86.6%; Pred. No. 5.4e-36;	AAY50690	
	Matches 97; Conservative 9; Mismatches 6; Indels 0; Gaps 0;	ID	AAY50690 standard; Protein; 134 AA.
		XX	
QY	1 DIVMTQSPDLSAVSLGERVTMCKSSQSVLYSSNQKNYLAWYQKPGQSPKLLIYWASTR 60	AC	AAY50690;
		XX	
Db	141 DIVLSQSPSSLAVSVGKVTMCKSSQSVLYSSNQKNYLAWYQKPGQSPKLLIYWASTR 200	DT	09-FEB-2000 (first entry)
		XX	
QY	61 ESGVPDRFSGSGGTDTFTLTISVQAEDLAVYYCHQYLSS-YTFGGGTTKLEIK 112	DE	Human Hum4 VL ClaI-HindIII segment encoded protein.
		XX	
Db	201 ESGVPDRFTGSGGTDTFTLTINSVKAEDLAVYYCQYYSTVTFEGAGTTKLEIK 252	KW	Human; antibody; humanized; anti-tumor; sialylated glycoprotein antibody;
		KW	TAG-72; cytostatic; cancer antigen; detection; carcinoma lesion;
		XX	diagnostic; treatment.
		OS	Homo sapiens.
		XX	
		PN	US5976531-A.
		XX	
		PD	02-NOV-1999.
		XX	
		PF	16-JUN-1994; 94US-0261354.
		XX	
		PR	19-APR-1990; 90US-0510697.
		PR	20-OCT-1992; 92US-0964536.
		XX	
		PA	(DOWC) DOW CHEM CO.
		XX	
		PI	Johnson KS, Mezes PS, Richard RA;
		XX	
		DR	WPI; 1999-632731/54.
		DR	N-PSDB; AAZ23970.
		XX	
		PT	New humanized anti-TAG-72 antibodies, used for the detection, in vivo
		PT	imaging and treatment of cancers -

xx Example 1; Figure 10A-E; 83pp; English.
PS
xx
CC This invention describes novel humanized anti-tumor associated
CC sialylated glycoprotein antibodies (TAG-72) which have cytostatic
CC activity. The antibodies have binding specificity for the cancer antigen
CC TAG-72. They can be used for the in vivo detection of carcinoma lesions.
CC They can also be used for in vitro diagnostics. They can also be modified
CC with therapeutic agents e.g. a radionuclide, drug, biological response
CC modifier, toxin or another antibody for the treatment of cancers. The
CC humanized anti-TAG-72 antibodies can reduce harmful anti-mouse antibody
CC hypersensitivity reactions.
xx
SQ Sequence 134 AA;

Query Match 87.9%; Score 512.5; DB 20; Length 134;
Best Local Similarity 88.5%; Pred. No. 3.1e-36;
Matches 100; Conservative 4; Mismatches 8; Indels 1; Gaps 1;
1 DIVMTQSPDRLAVSLGERVTMCKSSQSVLYSSNKNYLAHYQQKPGQSPKLLIYWASTR 60
|||||
21 DIVMTQSPDRLAVSLGERATINCKSSQSVLYSSNKNYLAHYQQKPGQSPKLLIYWASTR 80
|||||
61 ESGVDPDRFSGSGGTDTLTISVQAEDLAVYCHQYLS-SYTFGGGTKLEIK 112
|||||
81 ESGVDPDRFSGSGGTDTLTISVQAEDLAVYCHQYLS-SYTFGGGTKLEIK 133

Search completed: June 23, 2003, 14:01:39
Job time : 19.0709 secs

TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 342 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-828-741B-6

Query Match 89.3%; Score 520.5; DB 3; Length 342;
Best Local Similarity 89.4%; Pred. No. 1.7e-41;
Matches 101; Conservative 5; Mismatches 6; Indels 1; Gaps 1;
Qy 1 DIVMTQSPDLSAVSLGERVTMCKSSQSVLYSSNQKNYLAWYQKPGQSPKLLIYWASTR 60
Db 30 DIVMTQSPDLSAVSLGERATINCKSSQSVLYSSNSKNYLAWYQKPGQPPKLLIYWASTR 89
Db 61 ESGVPDRFSGSGGTDTFTLTSSVQAEDLAVYYCHQYLSS-YTFGGGTTKLEIK 112
Db 90 ESGVPDRFSGSGGTDTFTLTSSLQAEDVAVYYCQYVSTPYSFGQGTKEIK 142

RESULT 12
US-09-160-567-6
Sequence 6, Application US/09160567
Patent No. 6326179
GENERAL INFORMATION:
APPLICANT: Koentgen, Frank
APPLICANT: Suess, Gabriele M.
APPLICANT: Tarlington, David M.
APPLICANT: Treutlein, Herbert R.
TITLE OF INVENTION: CATALYTIC ANTIBODIES AND A METHOD OF
TITLE OF INVENTION: PRODUCING SAME
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
STREET: 400 Garden City Plaza
CITY: Garden City
STATE: New York
COUNTRY: United States of America
ZIP: 11530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/160,567
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/828,741
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Digiglio, Frank S.
REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 10591
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 342 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-160-567-6

Query Match 89.3%; Score 520.5; DB 4; Length 342;

Best Local Similarity 89.4%; Pred. No. 1.7e-41;
Matches 101; Conservative 5; Mismatches 6; Indels 1; Gaps 1;
Qy 1 DIVMTQSPDLSAVSLGERVTMCKSSQSVLYSSNQKNYLAWYQKPGQSPKLLIYWASTR 60
Db 30 DIVMTQSPDLSAVSLGERATINCKSSQSVLYSSNSKNYLAWYQKPGQPPKLLIYWASTR 89
Qy 61 ESGVPDRFSGSGGTDTFTLTSSVQAEDLAVYYCHQYLSS-YTFGGGTTKLEIK 112
Db 90 ESGVPDRFSGSGGTDTFTLTSSLQAEDVAVYYCQYVSTPYSFGQGTKEIK 142

RESULT 13
US-08-828-741B-4
Sequence 4, Application US/08828741B
Patent No. 6043069
GENERAL INFORMATION:
APPLICANT: Koentgen, Frank
APPLICANT: Suess, Gabriele M.
APPLICANT: Tarlington, David M.
APPLICANT: Treutlein, Herbert R.
TITLE OF INVENTION: CATALYTIC ANTIBODIES AND A METHOD OF
TITLE OF INVENTION: PRODUCING SAME
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
STREET: 400 Garden City Plaza
CITY: Garden City
STATE: New York
COUNTRY: United States of America
ZIP: 11530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/828,741B
FILING DATE: 26-MAR-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Digiglio, Frank S.
REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 10591
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 495 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-828-741B-4

Query Match 89.3%; Score 520.5; DB 3; Length 495;
Best Local Similarity 89.4%; Pred. No. 2.5e-41;
Matches 101; Conservative 5; Mismatches 6; Indels 1; Gaps 1;
Qy 1 DIVMTQSPDLSAVSLGERVTMCKSSQSVLYSSNQKNYLAWYQKPGQSPKLLIYWASTR 60
Db 30 DIVMTQSPDLSAVSLGERATINCKSSQSVLYSSNSKNYLAWYQKPGQPPKLLIYWASTR 89
Qy 61 ESGVPDRFSGSGGTDTFTLTSSVQAEDLAVYYCHQYLSS-YTFGGGTTKLEIK 112
Db 90 ESGVPDRFSGSGGTDTFTLTSSLQAEDVAVYYCQYVSTPYSFGQGTKEIK 142

RESULT 14
US-09-160-567-4
Sequence 4, Application US/09160567
Patent No. 6326179

GENERAL INFORMATION:
APPLICANT: Koentgen, Frank
APPLICANT: Suess, Gabriele M.
APPLICANT: Tarlinton, David M.
APPLICANT: Treutlein, Herbert R.
TITLE OF INVENTION: CATALYTIC ANTIBODIES AND A METHOD OF
TITLE OF INVENTION: PRODUCING SAME
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
STREET: 400 Garden City Plaza
CITY: Garden City
STATE: New York
COUNTRY: United States of America
ZIP: 11530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/160,567
FILING DATE:

CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/828,741
FILING DATE:

ATTORNEY/AGENT INFORMATION:
NAME: Digiglio, Frank S.
REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 10591
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366
TELEX: 230 901 SANS UR

INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 495 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-160-567-4

Query Match 89.3%; Score 520.5; DB 4; Length 495;
Best Local Similarity 89.4%; Pred. No. 2.5e-41;
Matches 101; Conservative 5; Mismatches 6; Indels 1; Gaps 1;

1 DIVMTQSPDSLAVSLGERVTMCKSSQSVLYSSNQKNYLAWYQKPGQSPKLLIYWASTR 60
|||||
30 DIVMTQSPDSLAVSLGERATINCKSSQSVLYSSNSKNYLAWYQKPGQPPKLLIYWASTR 89

QY 61 ESGVPDRFSGSGGTDTLTITSSVQAEDLAVYYCHQYLSS-YTFGGGTKLEIK 112
|||||
Db 90 ESGVPDRFSGSGGTDTLTITSSVQAEDVAVYYCQYYSTPYSPFGQGTKEIK 142

RESULT 15
US-07-916-098A-56
Sequence 56, Application US/07916098A
Patent No. 5871732
GENERAL INFORMATION:
APPLICANT: BURKLY, LINDA C.
APPLICANT: CHISHOLM, PATRICIA L.
APPLICANT: THOMAS, DAVID W.
APPLICANT: ROSA, MARGARET D.
APPLICANT: ROSA, JOSEPH J.
TITLE OF INVENTION: ANTI-CD4 ANTIBODY HOMOLOGS USEFUL IN
TITLE OF INVENTION: PROPHYLAXIS AND TREATMENT OF AIDS, ARC AND HIV INFECTION
NUMBER OF SEQUENCES: 61
CORRESPONDENCE ADDRESS:
ADDRESSEE: ALLEGRETTI & WITCOFF, LTD.
STREET: 10 SOUTH WACKER DRIVE
CITY: CHICAGO

STATE: ILLINOIS
COUNTRY: U.S.A.
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORD PERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/916,098A
FILING DATE: July 24, 1992
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US91/08843
FILING DATE: No. 5871732ember 27, 1991
CLASSIFICATION: 424
APPLICATION NUMBER: 07/618,542
FILING DATE: No. 5871732ember 27, 1990
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: JOHN J. MC DONNELL
REGISTRATION NUMBER: 26,949
REFERENCE/DOCKET NUMBER: 92,310-G
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 715-1000
TELEFAX: (312) 715-1234
TELEX: 910/221-5317
INFORMATION FOR SEQ ID NO: 56:
SEQUENCE CHARACTERISTICS:
LENGTH: 241 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-916-098A-56

Query Match 88.3%; Score 515; DB 2; Length 241;
Best Local Similarity 88.4%; Pred. No. 3.8e-41;
Matches 99; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

QY 1 DIVMTQSPDSLAVSLGERVTMCKSSQSVLYSSNQKNYLAWYQKPGQSPKLLIYWASTR 60
|||||
Db 23 DIVMTQSPDSLAVSLGERATINCKSSGSLLYSTNQKNYLAWYQKPGQPPKLLIYWASTR 82

QY 61 ESGVPDRFSGSGGTDTLTITSSVQAEDLAVYYCHQYLSSYTFGGGTKLEIK 112
|||||
Db 83 ESGVPDRFSGSGGTDTLTITSSVQAEDVAVYYCQYYSYRTFGRGKLEIK 134

RESULT 16
US-08-525-539A-80
Sequence 80, Application US/08525539A
Patent No. 6309636
GENERAL INFORMATION:
APPLICANT: DO COUTO, FERNANDO J.R.
APPLICANT: CERIANI, ROBERTO L.
APPLICANT: PETERSON, JERRY A.
TITLE OF INVENTION: RECOMBINANT PEPTIDES DERIVED FROM THE
TITLE OF INVENTION: MC3 ANTI-BA46 ANTIBODY, METHODS OF USE THEREOF, AND
TITLE OF INVENTION: METHODS OF HUMANIZING ANTIBODY PEPTIDES
NUMBER OF SEQUENCES: 81
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORRISON & FOERSTER
STREET: 755 Page Mill Road
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304-1018
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:

```
; APPLICATION NUMBER: US/08/525,539A
; FILING DATE: 14-SEP-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: DYLAN, TYLER
; REGISTRATION NUMBER: 37,612
; REFERENCE/DOCKET NUMBER: 27633-20001.21
; TELEPHONE: (415) 813-5600
; TELEFAX: (415) 494-0792
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 80:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 113 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-525-539A-80

Query Match      88.1%; Score 513.5; DB 4; Length 113;
Best Local Similarity 88.5%; Pred. No. 2.3e-41;
Matches 100; Conservative 5; Mismatches 7; Indels 1; Gaps 1;

Qy 1 DIVMTQSPDSLAVSLGERVTMCKSSQSVLYSSNQKNYLAWYQQKPGQSPKLLIYWASTR 60
   |||||
Db 1 DIVMTQSPDSLAVSLGERATINCKSSQSVLYSSNNKNYLAWYQQKPGQPPKLLIYWASTR 60
   |||||

Qy 61 ESGVPDRFSGSGGTDTFTLTSSVQAEDLAVYYCHQYLS-YTFGGGTKLEIK 112
   |||||
Db 61 ESGVPDRFSGSGGTDTFTLTSSLQAEDVAVYYCQYYSTPTFTFGGTVKVIK 113.

RESULT 17
US-08-463-903-4
; Sequence 4, Application US/08463903
; Patent No. 6071515
; GENERAL INFORMATION:
; APPLICANT: Mezes, Peter S.
; APPLICANT: Richard, Ruth A.
; APPLICANT: Affholter, Joseph A.
; APPLICANT: Kotite, Nicolas J.
; TITLE OF INVENTION: Dimer and Multimer Forms of Single Chain Polypeptides
; FILE REFERENCE: 40224A US
; CURRENT APPLICATION NUMBER: US/08/463,903
; CURRENT FILING DATE: 1995-06-05
; EARLIER APPLICATION NUMBER: US 07/935,695
; EARLIER FILING DATE: 1992-08-21
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: MS-Word for Windows, Ver. 7.0
; SEQ ID NO 4
; LENGTH: 133
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: Human kappa subgroup IV VL (Hum4 VL) from pRL1001
; LOCATION: 1..133
US-08-463-903-4

Query Match      87.9%; Score 512.5; DB 3; Length 133;
Best Local Similarity 88.5%; Pred. No. 3.4e-41;
Matches 100; Conservative 4; Mismatches 8; Indels 1; Gaps 1;

Qy 1 DIVMTQSPDSLAVSLGERVTMCKSSQSVLYSSNQKNYLAWYQQKPGQSPKLLIYWASTR 60
   |||||
Db 21 DIVMTQSPDSLAVSLGERATINCKSSQSVLYSSNNKNYLAWYQQKPGQPPKLLIYWASTR 80
   |||||

Qy 61 ESGVPDRFSGSGGTDTFTLTSSVQAEDLAVYYCHQYLS-SYTFGGGTKLEIK 112
   |||||
Db 81 ESGVPDRFSGSGGTDTFTLTSSLQAEDVAVYYCQYYSYPLTFGGGTVKVIK 133

RESULT 18
US-07-935-695-4
```

```
; Sequence 4, Application US/07935695
; Patent No. 6329507
; GENERAL INFORMATION:
; APPLICANT: Mezes, Peter S.
; APPLICANT: Richard, Ruth A.
; APPLICANT: Affholter, Joseph A.
; APPLICANT: Kotite, Nicolas J.
; TITLE OF INVENTION: Dimer and Multimer Forms of Single Chain Polypeptides
; FILE REFERENCE: 40224A US
; CURRENT APPLICATION NUMBER: US/07/935,695
; CURRENT FILING DATE: 1992-08-21
; PRIOR APPLICATION NUMBER: US 08/463,903
; PRIOR FILING DATE: 1995-06-05
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: MS-Word for Windows, Ver. 7.0
; SEQ ID NO 4
; LENGTH: 133
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: Human kappa subgroup IV VL (Hum4 VL) from pRL1001
; LOCATION: 1..133
US-07-935-695-4

Query Match      87.9%; Score 512.5; DB 4; Length 133;
Best Local Similarity 88.5%; Pred. No. 3.4e-41;
Matches 100; Conservative 4; Mismatches 8; Indels 1; Gaps 1;

Qy 1 DIVMTQSPDSLAVSLGERVTMCKSSQSVLYSSNQKNYLAWYQQKPGQSPKLLIYWASTR 60
   |||||
Db 21 DIVMTQSPDSLAVSLGERATINCKSSQSVLYSSNNKNYLAWYQQKPGQPPKLLIYWASTR 80
   |||||

Qy 61 ESGVPDRFSGSGGTDTFTLTSSVQAEDLAVYYCHQYLS-SYTFGGGTKLEIK 112
   |||||
Db 81 ESGVPDRFSGSGGTDTFTLTSSLQAEDVAVYYCQYYSYPLTFGGGTVKVIK 133

RESULT 19
US-08-463-903-6
; Sequence 6, Application US/08463903
; Patent No. 6071515
; GENERAL INFORMATION:
; APPLICANT: Mezes, Peter S.
; APPLICANT: Richard, Ruth A.
; APPLICANT: Affholter, Joseph A.
; APPLICANT: Kotite, Nicolas J.
; TITLE OF INVENTION: Dimer and Multimer Forms of Single Chain Polypeptides
; FILE REFERENCE: 40224A US
; CURRENT APPLICATION NUMBER: US/08/463,903
; CURRENT FILING DATE: 1995-06-05
; EARLIER APPLICATION NUMBER: US 07/935,695
; EARLIER FILING DATE: 1992-08-21
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: MS-Word for Windows, Ver. 7.0
; SEQ ID NO 6
; LENGTH: 275
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: Hum4VL-VH scFv from pSCFV UHH
; LOCATION: 1..275
US-08-463-903-6

Query Match      87.9%; Score 512.5; DB 3; Length 275;
Best Local Similarity 88.5%; Pred. No. 7.6e-41;
Matches 100; Conservative 4; Mismatches 8; Indels 1; Gaps 1;

Qy 1 DIVMTQSPDSLAVSLGERVTMCKSSQSVLYSSNQKNYLAWYQQKPGQSPKLLIYWASTR 60
   |||||
Db 23 DIVMTQSPDSLAVSLGERATINCKSSQSVLYSSNNKNYLAWYQQKPGQPPKLLIYWASTR 82
   |||||

Qy 61 ESGVPDRFSGSGGTDTFTLTSSVQAEDLAVYYCHQYLS-SYTFGGGTKLEIK 112
   |||||
```


APPLICATION NUMBER: EP 95 11 3021.0

QY 1 DIVMTQSPDSLAVSLGERVTMCKSSQSVLYSSNKNYLAQYQKPGQSPKLLIYWASTR 60
 :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
 Db 1 ELVMTQSPSSLAVSVEKVTGCKSSQSLLYSSNKNYLAQYQKPGQSPKLLIYWASTR 60
 QY 61 ESGVPDRFSGSGGTDFTLTISSVQAEDLAVYYCHQYLS-SYTFGGGTTKLEIK 112
 :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
 Db 61 ESGVPDRFTGSGGTDFTLTISSVKAEDLAVYYCQYYSYPTFTFGSGGTTKLEIK 113

 RESULT 23
 US-08-463-903-20
 ; Sequence 20, Application US/08463903
 ; Patent No. 6071515
 ; GENERAL INFORMATION:
 ; APPLICANT: Mezes, Peter S.
 ; APPLICANT: Richard, Ruth A.
 ; APPLICANT: Affholter, Joseph A.
 ; APPLICANT: Kotite, Nicolas J.
 ; TITLE OF INVENTION: Dimer and Multimer Forms of Single Chain Polypeptides
 ; FILE REFERENCE: 40224A US
 ; CURRENT APPLICATION NUMBER: US/08/463,903
 ; CURRENT FILING DATE: 1995-06-05
 ; EARLIER APPLICATION NUMBER: US 07/935,695
 ; EARLIER FILING DATE: 1992-08-21
 ; NUMBER OF SEQ ID NOS: 102
 ; SOFTWARE: MS-Word for Windows, Ver. 7.0
 ; SEQ ID NO 20
 ; LENGTH: 171
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; NAME/KEY: Hum4 VL-UNIHOPE linker-FLAG peptide construct in pATDFLAG
 ; LOCATION: 1..171
 US-08-463-903-20

 Query Match 87.6%; Score 510.5; DB 3; Length 171;
 Best Local Similarity 88.5%; Pred. No. 6.9e-41;
 Matches 100; Conservative 3; Mismatches 9; Indels 1; Gaps 1;

 QY 1 DIVMTQSPDSLAVSLGERVTMCKSSQSVLYSSNKNYLAQYQKPGQSPKLLIYWASTR 60
 :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
 Db 23 DIVMTQSPDSLAVSLGERATINCKSSQSVLYSSNKNYLAQYQKPGQSPKLLIYWASTR 82
 QY 61 ESGVPDRFSGSGGTDFTLTISSVQAEDLAVYYCHQYLS-SYTFGGGTTKLEIK 112
 :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
 Db 83 ESGVPDRFSGSGGTDFTLTISSVKAEDLAVYYCQYYSYPLTFTFGGTTKVVIK 135

 JULT 24
 US-07-935-695-20
 ; Sequence 20, Application US/07935695
 ; Patent No. 6329507
 ; GENERAL INFORMATION:
 ; APPLICANT: Mezes, Peter S.
 ; APPLICANT: Richard, Ruth A.
 ; APPLICANT: Affholter, Joseph A.
 ; APPLICANT: Kotite, Nicolas J.
 ; TITLE OF INVENTION: Dimer and Multimer Forms of Single Chain Polypeptides
 ; FILE REFERENCE: 40224A US
 ; CURRENT APPLICATION NUMBER: US/07/935,695
 ; CURRENT FILING DATE: 1992-08-21
 ; PRIOR APPLICATION NUMBER: US 08/463,903
 ; PRIOR FILING DATE: 1995-06-05
 ; NUMBER OF SEQ ID NOS: 102
 ; SOFTWARE: MS-Word for Windows, Ver. 7.0
 ; SEQ ID NO 20
 ; LENGTH: 171
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; NAME/KEY: Hum4 VL-UNIHOPE linker-FLAG peptide construct in pATDFLAG
 ; LOCATION: 1..171

DEVELOPMENTAL STAGE:
HAPLOTYPE:
TISSUE TYPE:
CELL TYPE: Hybridoma producing human
CELL TYPE: antibody GAH
CELL LINE:
ORGANELLE:
IMMEDIATE SOURCE:
LIBRARY:
CLONE:
POSITION IN GENOME:
CHROMOSOME/SEGMENT:
MAP POSITION:
UNITS:
FEATURE:
NAME/KEY:
LOCATION:
IDENTIFICATION METHOD:
OTHER INFORMATION:
PUBLICATION INFORMATION:
AUTHORS:
TITLE:
JOURNAL:
VOLUME:
ISSUE:
PAGES:
DATE:
DOCUMENT NUMBER:
FILING DATE:
PUBLICATION DATE:
RELEVANT RESIDUES IN SEQ ID NO:
US-08-360-125-6

Query Match 87.0%; Score 507.5; DB 1; Length 114;
Best Local Similarity 86.7%; Pred. No. 8.5e-41;
Matches 98; Conservative 7; Mismatches 7; Indels 1; Gaps 1;
Qy 1 DIVMTQSPDLSAVSLGERVTMCKSSQSVLYSSNOKNYLAWYQOKPGQSPKLLIYWASTR 60
Db 1 DIVMTQSPDLSAVSLGERATINCKSSQSVLYSNSNKKYLAHYQKPGQPPKLLIYWASTR 60
Qy 61 ESGVPDRFSGSGGTDFLTITSSVQAEDLAVYYCHQYLSS-YTFGGGTKLEIK 112
Db 61 ESGVPDRFSGSGGTDFLTITSSVQAEDLAVYYCHQYLSS-YTFGGGTKLEIK 113

Search completed: June 23, 2003, 14:04:52
time : 7.55319 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 23, 2003, 14:03:36 ; Search time 15.8865 Seconds
(without alignments)
762.858 Million cell updates/sec

Title: US-10-056-052A-18
Perfect score: 583
Sequence: 1 DIVMTQSPDSLAVSLGERVT.....YCHQYLSSYTFGGGTKLEIK 112

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 417779 seqs, 108206813 residues

1 number of hits satisfying chosen parameters: 417779

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database : Published_Applications_AA:*

1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*

2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*

3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*

4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*

5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*

6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*

7: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*

8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*

9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*

10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*

11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*

12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*

13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*

14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	583	100.0	112	9	US-10-056-052-18
2	553	94.9	112	9	US-10-056-052-10
3	545	93.5	112	9	US-10-056-052-6
4	544	93.3	112	9	US-10-056-052-14
5	527	90.4	113	9	US-09-894-839-2
6	527	90.4	113	9	US-09-988-013A-2
7	527	90.4	113	10	US-09-741-843-2
8	520.5	89.3	113	10	US-09-274-163E-16
9	516.5	88.6	114	10	US-09-274-163E-2
10	516.5	88.6	114	10	US-09-274-163E-6
11	516	88.5	130	9	US-10-146-305-7
12	515.5	88.4	114	10	US-09-274-163E-4
13	513.5	88.1	113	9	US-09-956-206A-80
14	512.5	87.9	240	10	US-09-799-514-8
15	510.5	87.6	113	9	US-09-995-529-6
16	507.5	86.0	114	10	US-09-810-502-38
17	503.5	86.4	113	9	US-10-121-464-6
18	503.5	86.4	114	9	US-10-125-687-11
19	503.5	86.4	114	12	US-10-025-687-11
					Sequence 18, Appl
					Sequence 10, Appl
					Sequence 6, Appl
					Sequence 14, Appl
					Sequence 2, Appl
					Sequence 2, Appl
					Sequence 2, Appl
					Sequence 16, Appl
					Sequence 2, Appl
					Sequence 6, Appl
					Sequence 4, Appl
					Sequence 8, Appl
					Sequence 6, Appl
					Sequence 38, Appl
					Sequence 6, Appl
					Sequence 11, Appl
					Sequence 11, Appl

20	502.5	86.2	135	9	US-10-171-452A-1	Sequence 1, Appl
21	500	85.8	113	9	US-09-894-839-6	Sequence 6, Appl
22	500	85.8	113	9	US-09-988-013A-6	Sequence 6, Appl
23	500	85.8	113	10	US-09-741-843-6	Sequence 6, Appl
24	499.5	85.7	113	9	US-10-121-464-2	Sequence 2, Appl
25	494	84.7	109	10	US-09-828-708-1	Sequence 1, Appl
26	494	84.7	109	10	US-09-828-708-4	Sequence 4, Appl
27	490.5	84.1	220	10	US-09-995-693-1	Sequence 1, Appl
28	490	84.0	132	9	US-09-249-011A-4	Sequence 4, Appl
29	487.5	83.6	113	9	US-10-144-644-21	Sequence 21, Appl
30	485.5	83.3	113	9	US-10-121-464-4	Sequence 4, Appl
31	482	82.7	239	9	US-09-249-011A-22	Sequence 22, Appl
32	481	82.5	132	9	US-09-249-011A-8	Sequence 8, Appl
33	479	82.2	101	9	US-10-194-975-89	Sequence 89, Appl
34	479	82.2	101	9	US-10-125-687-25	Sequence 25, Appl
35	479	82.2	101	12	US-10-025-687-25	Sequence 25, Appl
36	479	82.2	113	9	US-09-894-839-20	Sequence 20, Appl
37	477.5	81.9	154	9	US-09-925-299-1226	Sequence 1226, Ap
38	477.5	81.9	154	10	US-09-925-299-1226	Sequence 1226, Ap
39	475.5	81.6	113	9	US-09-999-025-7	Sequence 7, Appl
40	475.5	81.6	113	9	US-09-999-025-13	Sequence 13, Appl
41	475.5	81.6	113	9	US-09-999-040-7	Sequence 7, Appl
42	475.5	81.6	113	9	US-09-999-040-13	Sequence 13, Appl
43	475.5	81.6	113	9	US-09-998-817-7	Sequence 7, Appl
44	475.5	81.6	113	9	US-09-998-817-13	Sequence 13, Appl
45	475.5	81.6	113	9	US-09-999-021-7	Sequence 7, Appl
46	475.5	81.6	113	9	US-09-999-021-13	Sequence 13, Appl
47	475.5	81.6	113	9	US-10-040-997-7	Sequence 7, Appl
48	475.5	81.6	113	9	US-10-040-997-13	Sequence 13, Appl
49	475.5	81.6	444	9	US-10-013-173-49	Sequence 49, Appl
50	475.5	81.6	444	9	US-10-150-762-49	Sequence 49, Appl
51	474.5	81.4	153	9	US-09-187-693-62	Sequence 62, Appl
52	474.5	81.4	241	10	US-09-791-578-6	Sequence 6, Appl
53	474.5	81.4	241	10	US-09-791-540-6	Sequence 6, Appl
54	474.5	81.4	249	9	US-09-956-086-2	Sequence 2, Appl
55	474.5	81.4	249	9	US-09-956-087-2	Sequence 2, Appl
56	474.5	81.4	257	9	US-09-985-442-2	Sequence 2, Appl
57	474.5	81.4	257	10	US-09-791-578-4	Sequence 4, Appl
58	474.5	81.4	257	10	US-09-791-540-4	Sequence 4, Appl
59	474.5	81.4	257	10	US-09-983-580-2	Sequence 2, Appl
60	474.5	81.4	269	9	US-09-985-442-4	Sequence 4, Appl
61	474.5	81.4	269	10	US-09-983-580-4	Sequence 4, Appl
62	471.5	80.9	113	9	US-09-995-529-2	Sequence 2, Appl
63	469.5	80.5	246	10	US-09-791-578-2	Sequence 2, Appl
64	469.5	80.5	246	10	US-09-791-540-2	Sequence 2, Appl
65	468.5	80.4	113	9	US-10-144-644-46	Sequence 46, Appl
66	467.5	80.2	113	9	US-10-144-644-45	Sequence 45, Appl
67	462.5	79.3	113	9	US-10-144-644-16	Sequence 16, Appl
68	453	77.7	94	10	US-09-905-243-60	Sequence 60, Appl
69	452.5	77.6	115	10	US-09-998-831-9	Sequence 9, Appl
70	451.5	77.4	239	10	US-09-825-012-9	Sequence 9, Appl
71	445.5	76.4	220	10	US-09-917-410-5	Sequence 5, Appl
72	444.5	76.2	151	9	US-09-187-693-64	Sequence 64, Appl
73	443.5	76.1	113	9	US-09-999-025-9	Sequence 9, Appl
74	443.5	76.1	113	9	US-09-999-025-14	Sequence 14, Appl
75	443.5	76.1	113	9	US-09-999-040-9	Sequence 9, Appl
76	443.5	76.1	113	9	US-09-999-040-14	Sequence 14, Appl
77	443.5	76.1	113	9	US-09-998-817-9	Sequence 9, Appl
78	443.5	76.1	113	9	US-09-998-817-14	Sequence 14, Appl
79	443.5	76.1	113	9	US-09-999-021-9	Sequence 9, Appl
80	443.5	76.1	113	9	US-09-999-021-14	Sequence 14, Appl
81	443.5	76.1	113	9	US-10-040-997-9	Sequence 9, Appl
82	443.5	76.1	113	9	US-10-040-997-14	Sequence 14, Appl
83	441.5	75.7	113	9	US-10-144-644-70	Sequence 70, Appl
84	431.5	74.0	218	9	US-10-171-452A-39	Sequence 39, Appl
85	431.5	74.0	218	9	US-10-171-452A-45	Sequence 45, Appl
86	431.5	74.0	218	9	US-10-171-452A-51	Sequence 51, Appl
87	431.5	74.0	218	9	US-10-171-452A-57	Sequence 57, Appl
88	431.5	74.0	238	9	US-10-171-452A-38	Sequence 38, Appl
89	431.5	74.0	238	9	US-10-171-452A-44	Sequence 44, Appl
90	431.5	74.0	238	9	US-10-171-452A-50	Sequence 50, Appl
91	431.5	74.0	238	9	US-10-171-452A-56	Sequence 56, Appl
92	429.5	73.7	107	9	US-09-874-141-2	Sequence 2, Appl

93 429.5 73.7 133 9 US-09-874-141-47 Sequence 47, Appl
94 427.5 73.3 107 9 US-09-874-141-4 Sequence 4, Appli
95 425.5 73.0 151 9 US-09-187-693-58 Sequence 58, Appl
96 425.5 73.0 159 9 US-09-187-693-66 Sequence 66, Appl
97 424.5 72.8 131 9 US-09-879-461-58 Sequence 58, Appl
98 419.5 72.0 107 9 US-09-956-206A-81 Sequence 81, Appl
99 419.5 72.0 111 9 US-10-078-958-16 Sequence 16, Appl
100 419.5 72.0 127 9 US-09-956-206A-65 Sequence 65, Appl

ALIGNMENTS

RESULT 1
US-10-056-052-18
; Sequence 18, Application US/10056052
; Publication No. US20030099656A1
; GENERAL INFORMATION:
; APPLICANT: PATTI, Joseph M
; APPLICANT: HUTCHINS, Jeff T
; APPLICANT: DOMANSKI, Paul
; APPLICANT: PATEL, Pratiksha
; APPLICANT: HALL, Andrea
; TITLE OF INVENTION: MONOCLONAL ANTIBODIES TO THE CLFA PROTEIN . . .
; FILE REFERENCE: P07069US04/BAS
; CURRENT APPLICATION NUMBER: US/10/056,052
; CURRENT FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 60/308,116
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 60/298,413
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: 60/274,611
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: 60/264,072
; PRIOR FILING DATE: 2001-01-26
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 18
; LENGTH: 112
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-10-056-052-18

Query Match 100.0%; Score 583; DB 9; Length 112;
Best Local Similarity 100.0%; Pred. No. 3.1e-37;
Matches 112; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 DIVMTQSPDLSAVSLGERVTMCKSSQSVLYSSNQKNYLAWYQQKPGQSPKLLIYWASTR 60
|||||
1 DIVMTQSPDLSAVSLGERVTMCKSSQSVLYSSNQKNYLAWYQQKPGQSPKLLIYWASTR 60

61 ESGVPDRFSGSGGTDTLTITSSVQAEDLAVYYCHQYLSSYTFGGGTKLEIK 112
|||||
61 ESGVPDRFSGSGGTDTLTITSSVQAEDLAVYYCHQYLSSYTFGGGTKLEIK 112

RESULT 2
US-10-056-052-10
; Sequence 10, Application US/10056052
; Publication No. US20030099656A1
; GENERAL INFORMATION:
; APPLICANT: PATTI, Joseph M
; APPLICANT: HUTCHINS, Jeff T
; APPLICANT: DOMANSKI, Paul
; APPLICANT: PATEL, Pratiksha
; APPLICANT: HALL, Andrea
; TITLE OF INVENTION: MONOCLONAL ANTIBODIES TO THE CLFA PROTEIN . . .
; FILE REFERENCE: P07069US04/BAS
; CURRENT APPLICATION NUMBER: US/10/056,052
; CURRENT FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 60/308,116
; PRIOR FILING DATE: 2001-07-30
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 112
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-10-056-052-6

Query Match 93.5%; Score 545; DB 9; Length 112;
Best Local Similarity 92.0%; Pred. No. 2.2e-34;
Matches 103; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

1 DIVMTQSPDLSAVSLGERVTMCKSSQSVLYSSNQKNYLAWYQQKPGQSPKLLIYWASTR 60
|||||
1 NIMMTQSPSSLAVSAGEKVTMCKSSQSVLYSSNQKNYLAWYQQKPGQSPKLLIYWASTR 60

61 ESGVPDRFSGSGGTDTLTITSSVQAEDLAVYYCHQYLSSYTFGGGTKLEIK 112
|||||
61 ESGVPDRFSGSGGTDTLTITSSVQAEDLAVYYCHQYLSSYTFGGGTKLEIK 112

; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: 60/274,611
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: 60/264,072
; PRIOR FILING DATE: 2001-01-26
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10
; LENGTH: 112
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-10-056-052-10

Query Match 94.9%; Score 553; DB 9; Length 112;
Best Local Similarity 93.8%; Pred. No. 5.5e-35;
Matches 105; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 1 DIVMTQSPDLSAVSLGERVTMCKSSQSVLYSSNQKNYLAWYQQKPGQSPKLLIYWASTR 60
:|||||
Db 1 NIMMTQSPSSLAVSAGEKVTMCKSSQSVLYSSNQKNYLAWYQQKPGQSPKLLIYWASTR 60

QY 61 ESGVPDRFSGSGGTDTLTITSSVQAEDLAVYYCHQYLSSYTFGGGTKLEIK 112
|||||
Db 61 ESGVPDRFSGSGGTDTLTITSSVQAEDLAVYYCHQYLSSYTFGGGTKLEIK 112

RESULT 3

US-10-056-052-6
; Sequence 6, Application US/10056052
; Publication No. US20030099656A1
; GENERAL INFORMATION:
; APPLICANT: PATTI, Joseph M
; APPLICANT: HUTCHINS, Jeff T
; APPLICANT: DOMANSKI, Paul
; APPLICANT: PATEL, Pratiksha
; APPLICANT: HALL, Andrea
; TITLE OF INVENTION: MONOCLONAL ANTIBODIES TO THE CLFA PROTEIN . . .
; FILE REFERENCE: P07069US04/BAS
; CURRENT APPLICATION NUMBER: US/10/056,052
; CURRENT FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 60/308,116
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 60/298,413
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: 60/274,611
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: 60/264,072
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 112
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-10-056-052-6

Query Match 93.5%; Score 545; DB 9; Length 112;
Best Local Similarity 92.0%; Pred. No. 2.2e-34;
Matches 103; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

QY 1 DIVMTQSPDLSAVSLGERVTMCKSSQSVLYSSNQKNYLAWYQQKPGQSPKLLIYWASTR 60
:|||||
Db 1 NIMMTQSPSSLAVSAGEKVTMCKSSQSVLYSSNQKNYLAWYQQKPGQSPKLLIYWASTR 60

QY 61 ESGVPDRFSGSGGTDTLTITSSVQAEDLAVYYCHQYLSSYTFGGGTKLEIK 112
|||||
Db 61 ESGVPDRFSGSGGTDTLTITSSVQAEDLAVYYCHQYLSSYTFGGGTKLEIK 112

RESULT 4

US-10-056-052-14
; Sequence 14, Application US/10056052
; Publication No. US20030099656A1

GENERAL INFORMATION:
APPLICANT: PATTI, Joseph M
APPLICANT: HUTCHINS, Jeff T
APPLICANT: DOMANSKI, Paul
APPLICANT: PATEL, Pratiksha
APPLICANT: HALL, Andrea
TITLE OF INVENTION: MONOCLONAL ANTIBODIES TO THE CLFA PROTEIN
FILE REFERENCE: P07069US04/BAS
CURRENT APPLICATION NUMBER: US/10/056,052
CURRENT FILING DATE: 2002-04-19
PRIOR APPLICATION NUMBER: 60/308,116
PRIOR FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: 60/298,413
PRIOR FILING DATE: 2001-06-18
PRIOR APPLICATION NUMBER: 60/274,611
PRIOR FILING DATE: 2001-03-12
PRIOR APPLICATION NUMBER: 60/264,072
PRIOR FILING DATE: 2001-01-26
NUMBER OF SEQ ID NOS: 20
SOFTWARE: PatentIn version 3.1
SEQ ID NO 14
LENGTH: 112
TYPE: PRT
ORGANISM: Staphylococcus aureus
US-10-056-052-14

Query Match 93.3%; Score 544; DB 9; Length 112;
Best Local Similarity 92.0%; Pred. No. 2.6e-34;
Matches 103; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

QY 1 DIVMTQSPDSLAVSLGERVTMCKSSQSVLYSSNQKNYLAHYQQKPGQSPKLLIYWASTR 60
Db 1 NIMMTQSPSSLAVSAGEKVTMCRSSQSVLYSSNQKNYLAHYQQKPGQSPKLLIYWASTR 60

QY 61 ESGVPDRFSGSGTDFTLTISVQAEDLAVYCHQYLSSYTFGGGKLEIK 112
Db 61 ESGVPDRFSGSGTDFTLTISVQAEDLAVYCHQYLSSYTFGGGKLEIK 112

RESULT 5
US-09-894-839-2
Sequence 2, Application US/09894839
Publication No. US20030035800A1
GENERAL INFORMATION:
APPLICANT: LEUNG, Shui-on
APPLICANT: HANSEN, Hans
TITLE OF INVENTION: GLYCOSYLATED HUMANIZED B-CELL SPECIFIC ANTIBODIES
FILE REFERENCE: 018733/1049
CURRENT APPLICATION NUMBER: US/09/894,839
CURRENT FILING DATE: 2001-06-29
PRIOR APPLICATION NUMBER: US 09/155,107
PRIOR FILING DATE: 1998-11-17
PRIOR APPLICATION NUMBER: US 60/013,709
PRIOR FILING DATE: 1996-03-20
NUMBER OF SEQ ID NOS: 47
SOFTWARE: PatentIn version 3.1
SEQ ID NO 2
LENGTH: 113
TYPE: PRT
ORGANISM: Murinae gen. sp.
US-09-894-839-2

Query Match 90.4%; Score 527; DB 9; Length 113;
Best Local Similarity 88.4%; Pred. No. 5e-33;
Matches 99; Conservative 6; Mismatches 7; Indels 0; Gaps 0;

QY 1 DIVMTQSPDSLAVSLGERVTMCKSSQSVLYSSNQKNYLAHYQQKPGQSPKLLIYWASTR 60
Db 1 DIQLTQSPSSLAVSAGENVTMCKSSQSVLYSSNQKNYLAHYQQKPGQSPKLLIYWASTR 60

QY 61 ESGVPDRFSGSGTDFTLTISVQAEDLAVYCHQYLSSYTFGGGKLEIK 112

Db 61 ESGVPDRFSGSGTDFTLTISRQVEDLAIYCHQYLSSWTFGGGKLEIK 112

RESULT 6
US-09-988-013A-2
Sequence 2, Application US/09988013A
Publication No. US20030103979A1
GENERAL INFORMATION:
APPLICANT: LEUNG, Shui-on
APPLICANT: HANSEN, Hans
TITLE OF INVENTION: IMMUNOCONJUGATES AND HUMANIZED ANTIBODIES SPECIFIC FOR B-
FILE REFERENCE: 18733/1082
CURRENT APPLICATION NUMBER: US/09/988,013A
CURRENT FILING DATE: 2002-10-29
PRIOR APPLICATION NUMBER: US 09/741,843
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: US 09/127,902
PRIOR FILING DATE: 1998-08-03
PRIOR APPLICATION NUMBER: US 08/690,102
PRIOR FILING DATE: 1996-07-06
PRIOR APPLICATION NUMBER: US 08/289,576
PRIOR FILING DATE: 1994-08-12
NUMBER OF SEQ ID NOS: 25
SOFTWARE: PatentIn version 3.1
SEQ ID NO 2
LENGTH: 113
TYPE: PRT
ORGANISM: Murinae gen. sp.
US-09-988-013A-2

Query Match 90.4%; Score 527; DB 9; Length 113;
Best Local Similarity 88.4%; Pred. No. 5e-33;
Matches 99; Conservative 6; Mismatches 7; Indels 0; Gaps 0;

QY 1 DIVMTQSPDSLAVSLGERVTMCKSSQSVLYSSNQKNYLAHYQQKPGQSPKLLIYWASTR 60
Db 1 DIQLTQSPSSLAVSAGENVTMCKSSQSVLYSSNQKNYLAHYQQKPGQSPKLLIYWASTR 60

QY 61 ESGVPDRFSGSGTDFTLTISRQVEDLAIYCHQYLSSWTFGGGKLEIK 112
Db 61 ESGVPDRFSGSGTDFTLTISRQVEDLAIYCHQYLSSWTFGGGKLEIK 112

RESULT 7
US-09-741-843-2
Sequence 2, Application US/09741843
Publication No. US20020102254A1
GENERAL INFORMATION:
APPLICANT: LEUNG, Shui-on
APPLICANT: HANSEN, Hans
TITLE OF INVENTION: IMMUNOCONJUGATES AND HUMANIZED ANTIBODIES SPECIFIC FOR B-CELL
FILE REFERENCE: 018733/0996
CURRENT APPLICATION NUMBER: US/09/741,843
CURRENT FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: US 09/127,902
PRIOR FILING DATE: 1998-08-03
PRIOR APPLICATION NUMBER: US 08/690,102
PRIOR FILING DATE: 1996-07-06
PRIOR APPLICATION NUMBER: US 08/289,576
PRIOR FILING DATE: 1994-08-12
NUMBER OF SEQ ID NOS: 21
SOFTWARE: PatentIn version 3.1
SEQ ID NO 2
LENGTH: 113
TYPE: PRT
ORGANISM: Murinae gen. sp.
US-09-741-843-2

Query Match 90.4%; Score 527; DB 10; Length 113;
Best Local Similarity 88.4%; Pred. No. 5e-33;
Matches 99; Conservative 6; Mismatches 7; Indels 0; Gaps 0;


```
Qy 61 ESGVPDRFSGSGTDTLTISVQAEDLAVYYCHOYLSSYTFGGGKLEIK 112
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 61 ESGVPDRFTGSGSGTDTLTISVKAEDLAVYYCQYYSYRTFEGGKLEIK 112

RESULT 12
US-09-274-163E-4
; Sequence 4, Application US/09274163E
; Patent No. US20020137897A1
; GENERAL INFORMATION:
; APPLICANT: STEVENS, Fred J.
; APPLICANT: WILKINS STEVENS, Priscilla
; APPLICANT: RAFFEN, Rosemarie
; APPLICANT: SCHIFFER, Marianne
; TITLE OF INVENTION: OPTIMUM RECOMBINANT ANTIBODY PRODUCTION
; FILE REFERENCE: 051583/0224
; CURRENT APPLICATION NUMBER: US/09/274,163E
; CURRENT FILING DATE: 1999-03-22
; PRIOR APPLICATION NUMBER: US 08/373,380
; PRIOR FILING DATE: 1995-01-17
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 114
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-274-163E-4

Query Match 88.4%; Score 515.5; DB 10; Length 114;
Best Local Similarity 88.5%; Pred. No. 3.7e-32;
Matches 100; Conservative 5; Mismatches 7; Indels 1; Gaps 1;

Qy 1 DIVMTQSPDSLAVSLGERVTMNCSSQSVLYSSNQKNYLAWYQKPGQSPKLLIYWASTR 60
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1 DIVMTQSPDSLAVSLGERATINCKSSQSVLYSSNSKNYLAWYQDKPGQPPKLLIYWASTR 60

Qy 61 ESGVPDRFSGSGTDTLTISVQAEDLAVYYCHOYLSS-YTFGGGKLEIK 112
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 61 ESGVPDRFSGSGTDTLTISLQAEDVAVYYCQYYSTPYSFQGKLEIK 113

RESULT 13
US-09-956-206A-80
; Sequence 80, Application US/09956206A
; Patent No. US20020164339A1
; GENERAL INFORMATION:
; APPLICANT: DO COUTO, FERNANDO J.R.
; CERIANI, ROBERTO L.
; PETERSON, JERRY A.
; TITLE OF INVENTION: RECOMBINANT PEPTIDES DERIVED FROM THE
; MC3 ANTI-BA46 ANTIBODY, METHODS OF USE THEREOF, AND
; METHODS OF HUMANIZING ANTIBODY PEPTIDES
; NUMBER OF SEQUENCES: 81
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 755 Page Mill Road
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/956,206A
; FILING DATE: 19-Apr-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/525,539
; FILING DATE: 14-SEP-1995
```

```
; APPLICATION NUMBER: PCT/US95/11683
; FILING DATE: 14-SEP-1995
; APPLICATION NUMBER: 08/487,598
; FILING DATE: 7-JUNE-1995
; APPLICATION NUMBER: 08/307,868
; FILING DATE: 16-SEPT-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: WITT, ERIC
; REGISTRATION NUMBER: 44,408
; REFERENCE/DOCKET NUMBER: 276332000101
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 813-5600
; TELEFAX: (650) 494-0792
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 80:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 113 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 80:
US-09-956-206A-80

Query Match 88.1%; Score 513.5; DB 9; Length 113;
Best Local Similarity 88.5%; Pred. No. 5.2e-32;
Matches 100; Conservative 5; Mismatches 7; Indels 1; Gaps 1;

Qy 1 DIVMTQSPDSLAVSLGERVTMNCSSQSVLYSSNQKNYLAWYQKPGQSPKLLIYWASTR 60
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1 DIVMTQSPDSLAVSLGERATINCKSSQSVLYSSNNKNYLAWYQKPGQPPKLLIYWASTR 60

Qy 61 ESGVPDRFSGSGTDTLTISVQAEDLAVYYCHOYLSS-YTFGGGKLEIK 112
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 61 ESGVPDRFSGSGTDTLTISLQAEDVAVYYCQYYSTPTXFTFGQTKVEIK 113

RESULT 14
US-09-799-514-8
; Sequence 8, Application US/09799514
; Patent No. US20020065220A1
; GENERAL INFORMATION:
; APPLICANT: Young et al.
; TITLE OF INVENTION: Immunoglobulin Superfamily Polynucleotides, Polypeptides, and
; FILE REFERENCE: PT015P1
; CURRENT APPLICATION NUMBER: US/09/799,514
; CURRENT FILING DATE: 2001-03-07
; PRIOR APPLICATION NUMBER: PCT/US00/23662
; PRIOR FILING DATE: 2000-08-29
; PRIOR APPLICATION NUMBER: 60/152,248
; PRIOR FILING DATE: 1999-09-03
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 240
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-799-514-8

Query Match 87.9%; Score 512.5; DB 10; Length 240;
Best Local Similarity 87.6%; Pred. No. 1.2e-31;
Matches 99; Conservative 7; Mismatches 6; Indels 1; Gaps 1;

Qy 1 DIVMTQSPDSLAVSLGERVTMNCSSQSVLYSSNQKNYLAWYQKPGQSPKLLIYWASTR 60
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 21 DIVMTQSPDSLAVSLGERATINCKSSQTVLYSSDNKNYLAWYQKPGQPPKLLIYWASTR 80

Qy 61 ESGVPDRFSGSGTDTLTISVQAEDLAVYYCHOYLSS-YTFGGGKLEIK 112
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 81 ESGVPDRFSGSGTDTLTISLQAEDVAVYYCQYYSTPYSFQGKLEIK 133

RESULT 15
US-09-995-529-6
```



```
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 11
; LENGTH: 114
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Human consensus antibody light chain variable region
US-10-125-687-11

Query Match      86.4%; Score 503.5; DB 9; Length 114;
Best Local Similarity 86.7%; Pred. No. 2.9e-31;
Matches 98; Conservative 6; Mismatches 8; Indels 1; Gaps 1;

QY 1 DIVMTQSPDLSAVSLGERVTMCKSSQSVLYSSNKNYLAHYQQKPGQSPKLLIYWASTR 60
Db 1 DIVMTQSPDLSAVSLGERATINCRSSQSVLYSSNKNYLAHYQQKPGQSPKLLIYWASTR 60

QY 61 ESGVPDRFSGSGGTDFLTITSSVQAEDLAVYYCHQ-YLSSYTFGGGTTKLEIK 112
61 ESGVPDRFSGSGGTDFLTITSSVQAEDVAVYYCQOHYTPPTFGQGTKEIK 113

RESULT 19
US-10-025-687-11
; Sequence 11, Application US/10025687
; Patent No. US20020142255A1
; GENERAL INFORMATION:
; APPLICANT: Luo, Peter
; TITLE OF INVENTION: STRUCTURE-BASED CONSTRUCTION OF HUMAN ANTIBODY LIBRARY
; FILE REFERENCE: 26050-705
; CURRENT APPLICATION NUMBER: US/10/025,687
; CURRENT FILING DATE: 2002-04-17
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 11
; LENGTH: 114
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Human consensus antibody light chain variable region
US-10-025-687-11

Query Match      86.4%; Score 503.5; DB 12; Length 114;
Best Local Similarity 86.7%; Pred. No. 2.9e-31;
Matches 98; Conservative 6; Mismatches 8; Indels 1; Gaps 1;

QY 1 DIVMTQSPDLSAVSLGERVTMCKSSQSVLYSSNKNYLAHYQQKPGQSPKLLIYWASTR 60
1 DIVMTQSPDLSAVSLGERATINCRSSQSVLYSSNKNYLAHYQQKPGQSPKLLIYWASTR 60

QY 61 ESGVPDRFSGSGGTDFLTITSSVQAEDLAVYYCHQ-YLSSYTFGGGTTKLEIK 112
61 ESGVPDRFSGSGGTDFLTITSSVQAEDVAVYYCQOHYTPPTFGQGTKEIK 113

RESULT 20
US-10-171-452A-1
; Sequence 1, Application US/10171452A
; Publication No. US20030108518A1
; GENERAL INFORMATION:
; APPLICANT: Frewin, Mark
; APPLICANT: Waldmann, Herman
; APPLICANT: Gorman, Scott
; APPLICANT: Hale, Geoff
; APPLICANT: Rao, Patricia
; APPLICANT: Kornaga, Tadeusz
; APPLICANT: Ringler, Douglas
; APPLICANT: Cobbold, Stephen
; APPLICANT: Winsor-Hines, Dawn
; TITLE OF INVENTION: TRX1 Antibody and Uses Therefor
; FILE REFERENCE: 695458-59
; CURRENT APPLICATION NUMBER: US/10/171,452A
```

```
; CURRENT FILING DATE: 2003-02-10
; PRIOR APPLICATION NUMBER: US60/373,471
; PRIOR FILING DATE: 2002-04-18
; PRIOR APPLICATION NUMBER: US60/373,470
; PRIOR FILING DATE: 2002-04-18
; PRIOR APPLICATION NUMBER: US60/345,194
; PRIOR FILING DATE: 2002-10-19
; PRIOR APPLICATION NUMBER: GB0122724.8
; PRIOR FILING DATE: 2001-09-20
; PRIOR APPLICATION NUMBER: GB0114517.6
; PRIOR FILING DATE: 2001-06-14
; NUMBER OF SEQ ID NOS: 60
; SEQ ID NO 1
; LENGTH: 135
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-171-452A-1

Query Match      86.2%; Score 502.5; DB 9; Length 135;
Best Local Similarity 86.7%; Pred. No. 4e-31;
Matches 98; Conservative 6; Mismatches 8; Indels 1; Gaps 1;

QY 1 DIVMTQSPDLSAVSLGERVTMCKSSQSVLYSSNKNYLAHYQQKPGQSPKLLIYWASTR 60
Db 21 DIVMTQSPDLSAVSLGERATINCKSSQSVLYSSNKNYLAHYQQKPGQSPKLLIYWASTR 80

QY 61 ESGVPDRFSGSGGTDFLTITSSVQAEDLAVYYCHQYLSS-YTFGGGTTKLEIK 112
81 ESGVPDRFSGSGGTDFLTITSSVQAEDVAVYYCQYVSTPPMFGQGTKEIK 133

RESULT 21
US-09-894-839-6
; Sequence 6, Application US/09894839
; Publication No. US20030035800A1
; GENERAL INFORMATION:
; APPLICANT: LEUNG, Shui-On
; APPLICANT: HANSEN, Hans
; APPLICANT: QU, Zhengxing
; TITLE OF INVENTION: GLYCOSYLATED HUMANIZED B-CELL SPECIFIC ANTIBODIES
; FILE REFERENCE: 018733/1049
; CURRENT APPLICATION NUMBER: US/09/894,839
; CURRENT FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 09/155,107
; PRIOR FILING DATE: 1998-11-17
; PRIOR APPLICATION NUMBER: US 60/013,709
; PRIOR FILING DATE: 1996-03-20
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 113
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-894-839-6

Query Match      85.8%; Score 500; DB 9; Length 113;
Best Local Similarity 81.2%; Pred. No. 5.3e-31;
Matches 91; Conservative 13; Mismatches 8; Indels 0; Gaps 0;

QY 1 DIVMTQSPDLSAVSLGERVTMCKSSQSVLYSSNKNYLAHYQQKPGQSPKLLIYWASTR 60
Db 1 DIQLTQSPSSLSASVGDRTVMCKSSQSVLYSANHKNYLAHYQQKPGKAPKLLIYWASTR 60

QY 61 ESGVPDRFSGSGGTDFLTITSSVQAEDLAVYYCHQYLSSYTFGGGTTKLEIK 112
61 ESGVPSRFSGSGGTDFLTITSSVQAEDVAVYYCQYLSSTWTFGGGTTKVOIK 112

RESULT 22
US-09-988-013A-6
; Sequence 6, Application US/09988013A
; Publication No. US20030103979A1
; GENERAL INFORMATION:
```

```

; APPLICANT: LEUNG, Shui-on
; APPLICANT: HANSEN, Hans
; TITLE OF INVENTION: IMMUNOCONJUGATES AND HUMANIZED ANTIBODIES SPECIFIC FOR B-
; TITLE OF INVENTION: CELL LYMPHOMA AND LEUKEMIA CELLS
; FILE REFERENCE: 18733/1082
; CURRENT APPLICATION NUMBER: US/09/988,013A
; CURRENT FILING DATE: 2002-10-29
; PRIOR APPLICATION NUMBER: US 09/741,843
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/127,902
; PRIOR FILING DATE: 1998-08-03
; PRIOR APPLICATION NUMBER: US 08/690,102
; PRIOR FILING DATE: 1996-07-06
; PRIOR APPLICATION NUMBER: US 08/289,576
; PRIOR FILING DATE: 1994-08-12
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 113
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-988-013A-6

Query Match      85.8%; Score 500; DB 9; Length 113;
Best Local Similarity 81.2%; Pred. No. 5.3e-31;
Matches 91; Conservative 13; Mismatches 8; Indels 0; Gaps 0;

QY 1 DIVMTQSPDLSAVSLGERVTMCKSSQSVLYSSNQKNYLAWYQKPGQSPKLLIYWASTR 60
   ||:|||||||:|:|:|||||:|||||:|:|:|||||:|||||:|:|:|||||:|||||
Db 1 DIQLTQSPSSLSASVGDRTMCKSSQSVLYSANHKNYLAWYQKPGKAPKLLIYWASTR 60

QY 61 ESGVPDRFSGSGGTDTFTLTISVQAEDLAVYYCHQVLSSTYFGGGTKLEIK 112
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 61 ESGVPSRFSGSGGTDTFTLTISLQPEDIATYYCHQVLSSTYFGGGTKVQIK 112

RESULT 23
US-09-741-843-6
; Sequence 6, Application US/09741843
; Patent No. US20020102254A1
; GENERAL INFORMATION:
; APPLICANT: LEUNG, Shui-on
; APPLICANT: HANSEN, Hans
; TITLE OF INVENTION: IMMUNOCONJUGATES AND HUMANIZED ANTIBODIES SPECIFIC FOR B-CELL LYM
; TITLE OF INVENTION: AND LEUKEMIA CELLS
; FILE REFERENCE: 018733/0996
; CURRENT APPLICATION NUMBER: US/09/741,843
; CURRENT FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/127,902
; PRIOR FILING DATE: 1998-08-03
; PRIOR APPLICATION NUMBER: US 08/690,102
; PRIOR FILING DATE: 1996-07-06
; PRIOR APPLICATION NUMBER: US 08/289,576
; PRIOR FILING DATE: 1994-08-12
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 113
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-741-843-6

Query Match      85.8%; Score 500; DB 10; Length 113;
Best Local Similarity 81.2%; Pred. No. 5.3e-31;
Matches 91; Conservative 13; Mismatches 8; Indels 0; Gaps 0;

QY 1 DIVMTQSPDLSAVSLGERVTMCKSSQSVLYSSNQKNYLAWYQKPGQSPKLLIYWASTR 60
   ||:|||||||:|:|:|||||:|||||:|:|:|||||:|||||:|:|:|||||:|||||
Db 1 DIQLTQSPSSLSASVGDRTMCKSSQSVLYSANHKNYLAWYQKPGKAPKLLIYWASTR 60

QY 61 ESGVPDRFSGSGGTDTFTLTISVQAEDLAVYYCHQVLSSTYFGGGTKLEIK 112
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 61 ESGVPSRFSGSGGTDTFTLTISLQPEDIATYYCHQVLSSTYFGGGTKVQIK 112
```

```

RESULT 24
US-10-121-464-2
; Sequence 2, Application US/10121464
; Publication No. US20030103968A1
; GENERAL INFORMATION:
; APPLICANT: Boehringer Ingelheim International GmbH
; APPLICANT: Boehringer Ingelheim Pharmaceuticals, Inc.
; TITLE OF INVENTION: Cancer treatment by using FAP-alpha specific antibodies
; FILE REFERENCE: 1-1203ff
; CURRENT APPLICATION NUMBER: US/10/121,464
; CURRENT FILING DATE: 2002-11-15
; PRIOR APPLICATION NUMBER: US 60/283,868
; PRIOR FILING DATE: 2001-04-12
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 113
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Antibody
; OTHER INFORMATION: sequence
; US-10-121-464-2

Query Match      85.7%; Score 499.5; DB 9; Length 113;
Best Local Similarity 85.8%; Pred. No. 5.8e-31;
Matches 97; Conservative 6; Mismatches 9; Indels 1; Gaps 1;

QY 1 DIVMTQSPDLSAVSLGERVTMCKSSQSVLYSSNQKNYLAWYQKPGQSPKLLIYWASTR 60
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 DIVMTQSPDLSAVSLGERATINCKSSQSLLSYRNQKNYLAWYQKPGQPPKLLIFWASTR 60

QY 61 ESGVPDRFSGSGGTDTFTLTISVQAEDLAVYYCHQVLS-SYTFGGGTKLEIK 112
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 61 ESGVPDRFSGSGGTDTFTLTISLQAEDVAVYYCQYFSYPLTFGQGTKVEIK 113

RESULT 25
US-09-828-708-1
; Sequence 1, Application US/09828708
; Patent No. US20020146753A1
; GENERAL INFORMATION:
; APPLICANT: Ditzel, H.
; APPLICANT: Burton, D.
; APPLICANT: Schaller, M.
; TITLE OF INVENTION: Autoantibodies to glucose-6-phosphate isomerase and their part
; TITLE OF INVENTION: autoimmune disease
; FILE REFERENCE: 1361.005US1
; CURRENT APPLICATION NUMBER: US/09/828,708
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 123
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 109
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-828-708-1

Query Match      84.7%; Score 494; DB 10; Length 109;
Best Local Similarity 88.6%; Pred. No. 1.5e-30;
Matches 93; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

QY 8 PDSLAVSLGERVTMCKSSQSVLYSSNQKNYLAWYQKPGQSPKLLIYWASTRESCVPDR 67
   ||||| ||||| ||:|||||||:|:|:|||||:|:|:|||||:|:|:|||||:|:|:|||||
Db 1 PDSLAVSLGERATINCKSSQSVFYTSNNKNYLAWYQKPGQPPKLLIYWASTRESCVPDR 60

QY 68 FSGSGSGTDTFTLTISVQAEDLAVYYCHQVLSSTYFGGGTKLEIK 112
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 61 FSGSGSGTDTFTLTISLQAEDVAVYYCQYYDSYTFGQGTKLEIK 105
```


Search completed: June 23, 2003, 14:19:13
Job time : 16.8865 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run On: June 23, 2003, 13:59:31 ; Search time 7.5461 Seconds
(without alignments)
1426.837 Million cell updates/sec

Title: US-10-056-052A-18
Perfect score: 583
Sequence: 1 DIVMTQSPDSLAVSLGERVT.....YCHOYLSSYTFGGGTKLEIK 112

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues
number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database : PIR_73:*
1: pirl:*
2: pirl2:*
3: pirl3:*
4: pirl4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	530	90.9	111	2 G30502	Ig kappa chain v r
2	521	89.4	133	1 K4HUJI	Ig kappa chain pre
3	515.5	88.4	114	1 K4HULN	Ig kappa chain v-I
4	514.5	88.3	113	2 S34002	Ig kappa chain v r
5	514.5	88.3	134	2 S49531	anti-Sm antibody v
6	513.5	88.1	113	2 S34003	Ig kappa chain v r
7	512	87.8	112	2 S41393	Ig kappa chain v r
8	511	87.7	138	2 S26040	Ig kappa chain pre
9	510	87.5	112	2 S09970	Ig kappa chain v-J
10	507.5	87.0	113	2 S30523	Ig kappa chain v r
11	507.5	87.0	120	2 S51147	antibody light cha
12	507	87.0	103	2 PH1054	Ig light chain v r
13	502.5	86.2	113	2 S30520	Ig kappa chain v r
14	500	85.8	118	2 PT0356	Ig kappa chain v r
15	495.5	85.0	114	2 S44116	Ig kappa chain v-J
16	495.5	85.0	240	2 S06084	Ig kappa chain pre
17	494.5	84.8	114	2 S44119	Ig kappa chain v-J
18	494.5	84.8	145	2 PL0014	Ig kappa chain pre
19	492.5	84.5	134	1 K4HU17	Ig kappa chain pre
20	491	84.2	101	2 S26337	Ig light chain v r
21	490.5	84.1	132	2 S46373	Ig kappa chain v-J
22	489	83.9	112	2 S43103	Ig kappa chain v-J
23	489	83.9	113	2 PT0407	Ig light chain v r
24	486	83.4	138	2 A53261	Ig kappa chain pre
25	485.5	83.3	124	2 S40364	Ig kappa chain - h
26	484.5	83.1	134	2 PC1214	Ig kappa chain pre
27	484	83.0	133	2 PS0023	Ig kappa chain pre
28	483.5	82.9	129	2 S40347	Ig kappa chain - h
29	480.5	82.4	214	2 S68212	Ig kappa chain (Ma

30	479	82.2	121	1 K4HU	Ig kappa chain pre
31	475.5	81.6	113	2 A49260	antitumor monoclon
32	475	81.5	112	2 PL0265	Ig kappa chain v r
33	475	81.5	113	2 PT0408	Ig light chain v r
34	472.5	81.0	113	2 PL0263	Ig kappa chain v r
35	468.5	80.4	113	2 JC2270	PL7-6 antibody lig
36	466.5	80.0	109	2 S26336	Ig light chain v r
37	466	79.9	101	2 PH0869	Ig kappa chain v r
38	465.5	79.8	134	2 S21917	Ig kappa chain v r
39	464.5	79.7	135	2 S38807	Ig light chain v-J
40	463.5	79.5	220	2 A31790	Ig kappa chain v r
41	463	79.4	106	2 A49138	IgA kappa rheumato
42	462.5	79.3	112	2 E30538	Ig kappa chain v r
43	462.5	79.3	112	2 F30538	Ig kappa chain v r
44	456.5	78.3	113	2 PL0264	Ig kappa chain v r
45	453.5	77.8	136	2 A49137	Ig kappa chain pre
46	452	77.5	103	2 PH1047	Ig light chain v r
47	451.5	77.4	111	2 S03304	Ig kappa chain v r
48	451	77.4	104	2 PH1104	Ig light chain v r
49	449	77.0	104	2 PH1101	Ig light chain v r
50	448.5	76.9	96	2 G38601	Ig kappa chain v r
51	448	76.8	104	2 PH1102	Ig light chain v r
52	447	76.7	103	2 PH1052	Ig light chain v r
53	446.5	76.6	107	2 F30535	Ig kappa chain v r
54	446.5	76.6	108	2 E30535	Ig kappa chain v r
55	442	75.8	103	2 PH1050	Ig light chain v r
56	441.5	75.7	107	2 G30535	Ig kappa chain v r
57	441	75.6	104	2 PH1103	Ig light chain v r
58	440.5	75.6	129	2 S40329	Ig kappa chain v-J
59	440	75.5	102	2 A34153	Ig kappa chain v-I
60	439	75.3	103	2 PH1051	Ig light chain v r
61	438	75.1	101	2 PH1046	Ig light chain v r
62	438	75.1	120	2 G33932	Ig kappa chain pre
63	435.5	74.7	104	2 PH1053	Ig light chain v r
64	434.5	74.5	107	2 B30535	Ig kappa chain v r
65	434.5	74.5	107	2 D30535	Ig kappa chain v r
66	430	73.8	102	2 B34153	Ig kappa chain v-I
67	427.5	73.3	104	2 F38601	Ig kappa chain v r
68	427.5	73.3	105	2 C30535	Ig kappa chain v r
69	423.5	72.6	92	2 S37535	Ig kappa chain v r
70	423	72.6	98	2 S19974	Ig kappa chain v r
71	422.5	72.5	92	2 S37532	Ig kappa chain v r
72	419.5	72.0	216	2 JE0241	Ig kappa chain Am3
73	418.5	71.8	92	2 S37533	Ig kappa chain v r
74	412.5	70.8	92	2 S37529	Ig kappa chain v r
75	412	70.7	113	2 PT0409	Ig light chain v r
76	411.5	70.6	92	2 S37531	Ig kappa chain v r
77	410.5	70.4	92	2 S37530	Ig kappa chain v r
78	410.5	70.4	92	2 S37534	Ig kappa chain v r
79	407	69.8	94	2 S20648	Ig heavy chain v r
80	406	69.6	97	2 A42575	Ig kappa chain v r
81	405	69.5	225	2 S37484	Ig kappa chain - m
82	404.5	69.4	127	2 S04577	Ig kappa chain pre
83	402.5	69.0	107	2 S36268	Ig lambda chain v
84	401	68.8	107	2 S36275	Ig lambda chain v
85	399.5	68.5	117	2 S42466	Ig kappa chain v r
86	397	68.1	109	2 G30607	Ig kappa chain v-I
87	395	67.8	109	2 A30608	Ig kappa chain v-I
88	395	67.8	109	2 H30601	Ig kappa chain v-I
89	393.5	67.5	114	2 S46375	Ig kappa chain v-J
90	390	66.9	108	2 C30608	Ig kappa chain v-I
91	389.5	66.8	107	2 S32192	Ig kappa chain v r
92	388.5	66.6	108	2 H44151	Ig kappa chain v r
93	388.5	66.6	136	1 KVMS21	Ig kappa chain pre
94	388	66.6	124	2 S20633	Ig kappa chain - h
95	387.5	66.5	111	2 D37266	Ig kappa chain v r
96	387.5	66.5	117	2 S40362	Ig kappa chain - h
97	387	66.4	109	2 G30601	Ig kappa chain v-I
98	387	66.4	110	1 KVMS10	Ig kappa chain v r
99	386.5	66.3	130	2 C29380	Ig kappa chain pre
100	386.5	66.3	152	2 S30751	Ig kappa chain pre

Query Match 88.3%; Score 514.5; DB 2; Length 113;
 Best Local Similarity 88.5%; Pred. No. 1.2e-39;
 Matches 100; Conservative 4; Mismatches 8; Indels 1; Gaps 1;

QY 1 DIVMTQSPDSLAVSLGERVTMCKSSQSVLYSSNQKNYLAWYQQKPGQSPKLLIYWASTR 60
 |||||
 Db 1 DIVMTQSPDSLAVSLGERATINCKSSQSVLYSSNNKNYLAWYQQKPGQPPKLLIYWASTR 60
 |||||

QY 61 ESGVPDRFSGSGSGTDFTLTISVQAEDLAVYYCHQYLS-SYTFGGGTTKLEIK 112
 |||||
 Db 61 ESGVPDRFSGSGSGTDFTLTISLQAEDVAVYYCHQYIGIPRTFGGQTKVEIK 113
 |||||

RESULT 5
 S49531
 anti-Sm antibody VL chain (V kappa 4/J kappa 3) - human
 C:Species: Homo sapiens (man)
 C:Date: 01-Feb-1995 #sequence_revision 12-May-1995 #text_change 21-Jan-2000
 C:Accession: S49531
 R;Moudi, M.; Edwards, J.; Cairns, E.; Bell, D.
 A:Submitted to the EMBL Data Library, October 1994
 A:Description: Molecular characterization of natural human anti-Sm autoantibodies.
 A:Reference number: S48797
 A:Accession: S49531
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-134 <MAH>
 A:Cross-references: EMBL:Z46347; NID:g560841; PIDN:CAA86466.1; PID:g560842
 C:Superfamily: immunoglobulin V region; immunoglobulin homology
 F:36-116/Domain: immunoglobulin homology <IMM>

Query Match 88.3%; Score 514.5; DB 2; Length 134;
 Best Local Similarity 87.6%; Pred. No. 1.4e-39;
 Matches 99; Conservative 7; Mismatches 6; Indels 1; Gaps 1;

QY 1 DIVMTQSPDSLAVSLGERVTMCKSSQSVLYSSNQKNYLAWYQQKPGQSPKLLIYWASTR 60
 |||||
 Db 21 DIVMTQSPDSLAVSLGERATINCKSSQSVLYSSNNKNYLAWYQQKPGQPPKLLIYWASTR 80
 |||||

QY 61 ESGVPDRFSGSGSGTDFTLTISVQAEDLAVYYCHQYLS-SYTFGGGTTKLEIK 112
 |||||
 Db 81 ESGVPDRFSGSGSGTDFTLTISLQAEDVAVYYCQYISTAFTEGPGTKVDIK 133
 |||||

RESULT 6
 S34003
 Ig kappa chain V region - human
 C:Species: Homo sapiens (man)
 C:Date: 02-Dec-1993 #sequence_revision 10-Nov-1995 #text_change 21-Jan-2000
 C:Accession: S34003
 R;Marette, X.; Tsapis, A.; Brouet, J.C.
 Eur. J. Immunol. 23, 846-851, 1993
 A:Title: Nucleotide sequence analysis of the variable domains of four human monoclonal
 A:Reference number: S34001; MUID:93209281; PMID:7681398
 A:Accession: S34003
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-113 <MAR>
 C:Superfamily: immunoglobulin V region; immunoglobulin homology
 C:Keywords: heterotetramer; immunoglobulin
 F:16-96/Domain: immunoglobulin homology <IMM>

Query Match 88.1%; Score 513.5; DB 2; Length 113;
 Best Local Similarity 88.5%; Pred. No. 1.5e-39;
 Matches 100; Conservative 6; Mismatches 6; Indels 1; Gaps 1;

QY 1 DIVMTQSPDSLAVSLGERVTMCKSSQSVLYSSNQKNYLAWYQQKPGQSPKLLIYWASTR 60
 |||||
 Db 1 DIVMTQSPDSLAVSLGERATINCKSSQSVLYSSNNKNYLAWYQQKAGQPPKLLIYWASTR 60
 |||||

QY 61 ESGVPDRFSGSGSGTDFTLTISVQAEDLAVYYCHQYLS-SYTFGGGTTKLEIK 112
 |||||
 Db 61 ESGVPDRFSGSGSGTDFTLTISLQAEDVAVYYCQYLTPTPTFGGQTKVEIK 113
 |||||

Query Match 86.2%; Score 502.5; DB 2; Length 113;
Best Local Similarity 85.0%; Pred. No. 1.5e-38;
Matches 96; Conservative 8; Mismatches 8; Indels 1; Gaps 1;

QY 1 DIVMTQSPDLSAVSLGERVTMCKSSQSVLYSSNKNYLAHYQKPGQSPKLLIYWASTR 60
Db 1 DIVMTQSPDLSAVSLGERATINCKSSQSLFYSSNKNYLAHYQKPGQSPKLLIYWASIR 60
QY 61 ESGVPDRFSGSGGTDFTLTISSVQAEDLAVYCHQYLSS-YTFGGGKLEIK 112
Db 61 ESGVPDRFSGSGGTDFTLTISSLQAEDVAVYCOQYNTPLTFGGGKVEIK 113

RESULT 14
PT0356
Ig kappa chain V region (2B11.1) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 21-Jan-2000
C:Accession: PT0356
C:Keywords: R.; Kleiner, G.; Turken, A.; Papazian, L.; Diamond, B.
J. Exp. Med. 173, 287-296, 1991
A:Title: A novel class of anti-DNA antibodies identified in BALB/c mice.
A:Reference number: PT0352; MUID:91108325; PMID:1988536
A:Accession: PT0356
A:Molecule type: mRNA
A:Residues: 1-118 <SHE>
A:Experimental source: strain BALB/c
A:Note: the authors translated the codon CTT for residue 32 as Ser
C:Comment: This protein is an anti-double-stranded DNA antibody.
C:Superfamily: immunoglobulin V region; immunoglobulin homology.
F:19-99/Domain: immunoglobulin homology <IMM>

Query Match 85.8%; Score 500; DB 2; Length 118;
Best Local Similarity 86.6%; Pred. No. 2.6e-38;
Matches 97; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

QY 1 DIVMTQSPDLSAVSLGERVTMCKSSQSVLYSSNKNYLAHYQKPGQSPKLLIYWASTR 60
Db 4 DIVMSQSPSSLAVSVGEKVTMCKSSQSLLYSSNKNYLAHYQKPGQSPKLLIYWASTR 63
QY 61 ESGVPDRFSGSGGTDFTLTISSVQAEDLAVYCHQYLSSYTFGGGKLEIK 112
Db 64 ESGVPDRFTGSGSGGTDFTLTISSVKAEDLAVYCOQYYSYPTTFGGGKLEIK 115

RESULT 15
S44116
Ig kappa chain V-J region - human
C:Species: Homo sapiens (man)
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 24-May-2001
C:Accession: S44116
C:Keywords: R.; E.; Zhu, D.; Ovecka, M.; Winter, G.; Hamblin, T.J.; Stevenson, F.K.
submitted to the EMBL Data Library, March 1994
A:Description: Idiotype vaccination against human B-cell lymphoma: rescue of variable
A:Reference number: S44105
A:Accession: S44116
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-114 <HAW>
C:Cross-references: EMBL:Z31391; NID:g472970; PIDN:CAA83266.1; PID:g940527
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:16-96/Domain: immunoglobulin homology <IMM>

Query Match 85.0%; Score 495.5; DB 2; Length 114;
Best Local Similarity 85.0%; Pred. No. 6.3e-38;
Matches 96; Conservative 7; Mismatches 9; Indels 1; Gaps 1;

QY 1 DIVMTQSPDLSAVSLGERVTMCKSSQSVLYSSNKNYLAHYQKPGQSPKLLIYWASTR 60
Db 1 DIVMTQSPDLSVSLGERATINCKSSQSLLYSSNKNYLAHYQKPGQSPKLLIKWASTR 60

QY 61 ESGVPDRFSGSGGTDFTLTISSVQAEDLAVYCHQYLSS-YTFGGGKLEIK 112
Db 61 ESGVPDRFSGSGGTDFTLTISSLQAEDVAVYCOQYNTPLTFGGGKVEIK 113

RESULT 16
S06084
Ig kappa chain precursor - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 28-Feb-1990 #sequence_revision 28-Feb-1990 #text_change 21-Jan-2000
C:Accession: S06084
R:Crowe, J.S.; Smith, M.A.; Cooper, H.J.
Nucleic Acids Res. 17, 7992, 1989
A:Title: Nucleotide sequence of Y3-Ag 1.2.3. rat myeloma immunoglobulin kappa chain c
A:Reference number: S06084; MUID:90016888; PMID:2508067
A:Accession: S06084
A:Molecule type: mRNA
A:Residues: 1-240 <CRO>
A:Cross-references: EMBL:X16129; NID:g56457; PIDN:CAA34256.1; PID:g56458
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:1-20/Domain: signal sequence #status predicted <SIG>
F:21-240/Product: Ig kappa chain #status predicted <MAT>
F:153-222/Domain: immunoglobulin homology <IMM>

Query Match 85.0%; Score 495.5; DB 2; Length 240;
Best Local Similarity 84.1%; Pred. No. 1.4e-37;
Matches 95; Conservative 6; Mismatches 11; Indels 1; Gaps 1;

QY 1 DIVMTQSPDLSAVSLGERVTMCKSSQSVLYSSNKNYLAHYQKPGQSPKLLIYWASTR 60
Db 21 DFVMTQSPSSLAVSAGETVTINCKSSQSLFYSSNKNYLAHYQKPGQSPKLLIYWASTR 80

QY 61 ESGVPDRFSGSGGTDFTLTISSVQAEDLAVYCHQYLSS-YTFGGGKLEIK 112
Db 81 QSGVPDRFIGSGGTDFTLTISSVQAEDLAIYVCLQYVETPTTFGAGTKLEK 133

RESULT 17
S44119
Ig kappa chain V-J region - human
C:Species: Homo sapiens (man)
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 24-May-2001
C:Accession: S44119
R:Hawkins, R.E.; Zhu, D.; Ovecka, M.; Winter, G.; Hamblin, T.J.; Stevenson, F.K.
submitted to the EMBL Data Library, March 1994
A:Description: Idiotype vaccination against human B-cell lymphoma: rescue of variabl
A:Reference number: S44105
A:Accession: S44119
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-114 <HAW>
C:Cross-references: EMBL:Z31396; NID:g472973; PIDN:CAA83271.1; PID:g940530
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:16-96/Domain: immunoglobulin homology <IMM>

Query Match 84.8%; Score 494.5; DB 2; Length 114;
Best Local Similarity 85.0%; Pred. No. 7.7e-38;
Matches 96; Conservative 5; Mismatches 11; Indels 1; Gaps 1;

QY 1 DIVMTQSPDLSAVSLGERVTMCKSSQSVLYSSNKNYLAHYQKPGQSPKLLIYWASTR 60
Db 1 DIQMTQSPDSLTVSLGERATINCKSSQSLIYSSNDKNYLAHYQKAGQPPKLLIYWASTR 60

QY 61 ESGVPDRFSGSGGTDFTLTISSVQAEDLAVYCHQYLSS-YTFGGGKLEIK 112
Db 61 ESGVPDRFSGSGSATDFTLTISSLQAEDVAVYCOQYYSIPTLTFGGGKVEIK 113

RESULT 18
PL0014
Ig kappa chain precursor V region (F6-3) - mouse (fragment)

C;Species: Mus musculus (house mouse)
C;Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 21-Jan-2000
C;Accession: PL0014
R;Cheng, H.L.; Sood, A.K.; Ward, R.E.; Kieber-Emmons, T.; Kohler, H.
Mol. Immunol. 25, 33-40, 1988
A;Title: Structural basis of stimulatory anti-idiotypic antibodies.
A;Reference number: PL0011; MUID:88142863; PMID:3125424
A;Accession: PL0014
A;Molecule type: mRNA
A;Residues: 1-145 <CHE>
A;Experimental source: cell line F6-3
C;Comment: This protein is an anti-idiotypic antibody that induces an anti-phosphorylcholine
C;Superfamily: immunoglobulin V region; immunoglobulin homology
C;Keywords: heterotetramer; immunoglobulin
F;1-20/Domain: signal sequence #status predicted <SIG>
F;21-134/Product: Ig heavy chain V region (4C11) #status predicted <MAT>
F;36-116/Domain: immunoglobulin homology <IMM>
F;44-60/Region: complementarity-determining 1
F;76-82/Region: complementarity-determining 2
F;83-114/Region: complementarity-determining 3
F;122-134/Region: complementarity-determining 4
F;43-114/Disulfide bonds: #status predicted

Query Match 84.8%; Score 494.5; DB 2; Length 145;
Best Local Similarity 84.2%; Pred. No. 9.9e-38;
Matches 96; Conservative 11; Mismatches 4; Indels 3; Gaps 2;

QY 1 DIVMTQSPDLSAVSLGERVTMCKSSQSVLYSSNKNYLAQQKPGQSPKLLIYWASTR 60
Db 21 DIVMSQSPSSLAIVSGEKVTMCKSSQSVLYSSNKNYLAQQKPGQSPKLLIYWASTR 80

QY 61 ESGVPDRFSGSGTDTFTLTSSVQAEDLAVYYCHQYLS--TFGGGTKLEIK 112
Db 81 ESGVPDRFTGSGTDTFTLTSSVKAEDLAVYYCQY-DSYPLTFGSGTKLEMK 133

RESULT 19
K4HU17
Ig kappa chain precursor V-IV region (B17) - human
C;Species: Homo sapiens (man)
C;Date: 30-Jun-1987 #sequence_revision 30-Jun-1987 #text_change 21-Jan-2000
C;Accession: A01905
R;Marsh, P.; Mills, F.; Gould, H.
Nucleic Acids Res. 13, 6531-6544, 1985
A;Title: Detection of a unique human V kappa IV germline gene by a cloned cDNA probe.
A;Reference number: A01905; MUID:86041854; PMID:2997713
A;Accession: A01905
A;Molecule type: mRNA
A;Residues: 1-134 <MAR>
C;Comment: The sequence was determined from the differentiated gene
C;Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kappa) chain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into larger complexes.
C;Superfamily: immunoglobulin V region; immunoglobulin homology
C;Keywords: heterotetramer; immunoglobulin
F;1-20/Domain: signal sequence #status predicted <SIG>
F;21-134/Product: Ig kappa chain V-IV region (B17) #status predicted <MAT>
F;21-43/Region: framework 1
F;36-116/Domain: immunoglobulin homology <IMM>
F;44-60/Region: complementarity-determining 1
F;61-75/Region: framework 2
F;76-82/Region: complementarity-determining 2
F;83-114/Region: framework 3
F;115-121/Region: complementarity-determining 3
F;122-134/Region: framework 4
F;43-114/Disulfide bonds: #status predicted

Query Match 84.5%; Score 492.5; DB 1; Length 134;
Best Local Similarity 85.0%; Pred. No. 1.4e-37;
Matches 96; Conservative 8; Mismatches 8; Indels 1; Gaps 1;

QY 1 DIVMTQSPDLSAVSLGERVTMCKSSQSVLYSSNKNYLAQQKPGQSPKLLIYWASTR 60
Db 21 DIVMTQSPDLSAVSLGERATINCKSSQSVLYSSNKNYLAQQKPGQSPKLLIYCASTR 80

QY 61 ESGVPDRFSGSGTDTFTLTSSVQAEDLAVYYCHQYLS-SYTFGGGTKLEIK 112
Db 81 ESGVPDRFSGSGTDTFTLTSSVQAEDVAVYYCQYYNLPWTFGGQGTKVEIK 133

RESULT 20
S26337
Ig light chain V region - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 19-Mar-1998 #sequence_revision 19-Mar-1998 #text_change 20-Jun-2000
C;Accession: S26337; S78449
R;Stark, S.E.; Caton, A.J.
J. Exp. Med. 174, 613-624, 1991
A;Title: Antibodies that are specific for a single amino acid interchange in a protein
A;Reference number: S26309; MUID:91341421; PMID:1908510
A;Accession: S26337
A;Molecule type: mRNA
A;Residues: 1-101 <STA>
A;Cross-references: EMBL:X59193
R;Caton, A.J.
submitted to the EMBL Data Library, April 1991
A;Reference number: S78447
A;Accession: S78449
A;Molecule type: mRNA
A;Residues: 1-60, 'T', '62-91, 'S', '93-101 <CAT>
A;Cross-references: EMBL:X59193; NID:g52323; PIDN:CAA41903.1; PID:g1334067
C;Superfamily: immunoglobulin V region; immunoglobulin homology
C;Keywords: heterotetramer; immunoglobulin
F;8-88/Domain: immunoglobulin homology <IMM>

Query Match 84.2%; Score 491; DB 2; Length 101;
Best Local Similarity 93.0%; Pred. No. 1.4e-37;
Matches 93; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 10 SLAVSLGERVTMCKSSQSVLYSSNKNYLAQQKPGQSPKLLIYWASTRESGVDPDRFS 69
Db 2 SLAVSAGEKVTMCKSSQSVLYSSNKNYLAQQKPGQSPKLLIYWASTRESGVDPDRFI 61

QY 70 GSGSGTDTFTLTSSVQAEDLAVYYCHQYLSSTYTFGGGTKL 109
Db 62 GSGSGTDTFTLTSSVQAEDLAVYYCHQYLSSTYTFGGGTKL 101

RESULT 21
S46373
Ig kappa chain V-J region (T24-9) - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 27-Jan-1995 #sequence_revision 01-Sep-1995 #text_change 21-Jan-2000
C;Accession: S46373; S38647
R;Bensimon, C.; Chastagner, P.; Zouali, M.
EMBO J. 13, 2951-2962, 1994
A;Title: Human lupus anti-DNA autoantibodies undergo essentially primary V(chi) gene
A;Reference number: S46369; MUID:94313975; PMID:8039491
A;Accession: S46373
A;Molecule type: mRNA
A;Residues: 1-132 <BEN>
A;Cross-references: EMBL:Z27174; NID:g415963; PIDN:CAA81698.1; PID:g415964
C;Superfamily: immunoglobulin V region; immunoglobulin homology
C;Keywords: heterotetramer; immunoglobulin
F;34-114/Domain: immunoglobulin homology <IMM>

Query Match 84.1%; Score 490.5; DB 2; Length 132;
Best Local Similarity 84.1%; Pred. No. 2.1e-37;
Matches 95; Conservative 9; Mismatches 8; Indels 1; Gaps 1;

QY 1 DIVMTQSPDLSAVSLGERVTMCKSSQSVLYSSNKNYLAQQKPGQSPKLLIYWASTR 60
Db 19 DIVMTQSPDLSAVSLGERATINCKSSQSVLYSSNKNYLAQQKPGQSPKLLIHWASSR 78

QY 61 ESGVPDRFSGSGTDTFTLTSSVQAEDLAVYYCHQYLS--YTFGGGTKLEIK 112
Db 79 ESGVLDRFSGSGTDTFTLTSSVQAEDLAVYYCHQYSSPFTFGPGTKVDIK 131

RESULT 22
S43103
Ig kappa chain V-J region (4B1 VL) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 06-Feb-1995 #sequence_revision 06-Feb-1995 #text_change 24-May-2001
C;Accession: S43103
R;Gilbert, D.; Brard, F.; Margaritte, C.; Delpech, A.; Tron, F.
submitted to the EMBL Data Library, March 1994
A;Description: An Idiotypic D23-bearing polyclonal antibody
A;Reference number: S42484
A;Accession: S43103
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-112 <IGL>
A;Cross-references: EMBL:Z31353; NID:G467574; PIDN:CAA83231.1; PID:G467575
C;Superfamily: immunoglobulin V region; immunoglobulin homology
C;Keywords: heterotetramer; immunoglobulin
F;96/Domain: immunoglobulin homology <IMM>
Query Match 83.9%; Score 489; DB 2; Length 112;
Best Local Similarity 84.8%; Pred. No. 2.4e-37;
Matches 95; Conservative 7; Mismatches 10; Indels 0; Gaps 0;
QY 1 DIVMTQSPDLSAVSLGERVTMCKSSQSVLYSSNQKNYLAWYQQKPGQSPKLLIYWASTR 60
Db 1 DIVMSQSPSSLAWSAGEKVTMCKSSQSLNLSRTRKNYLAWYQQKPGQSPKLLIYWASTR 60
QY 61 ESGVPDRFSGSGGTDFLTITSSVQAEDLAVYYCHQYLSYTFGGGTTKLEIK 112
Db 61 ESGVPDRFTGSGGTDFLTITSSVQAEDLAVYYCKQSYNLRFTGGGTTKLEIK 112
RESULT 23
PT0407
Ig light chain V region (S107/VH11 group 1-4) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 21-Jan-2000
C;Accession: PT0407
R;Behar, S.M.; Lustgarten, D.L.; Corbet, S.; Scharff, M.D.
J. Exp. Med. 173, 731-741, 1991
A;Title: Characterization of somatically mutated S107 VH11-encoded anti-DNA autoantibody
A;Reference number: PT0376; MUID:91147903; PMID:1900082
A;Accession: PT0407
A;Molecule type: DNA
A;Residues: 1-113 <BEH>
C;Superfamily: immunoglobulin V region; immunoglobulin homology
C;Keywords: immunoglobulin
F;96/Domain: immunoglobulin homology <IMM>
Query Match 83.9%; Score 489; DB 2; Length 113;
Best Local Similarity 83.9%; Pred. No. 2.4e-37;
Matches 94; Conservative 8; Mismatches 10; Indels 0; Gaps 0;
QY 1 DIVMTQSPDLSAVSLGERVTMCKSSQSVLYSSNQKNYLAWYQQKPGQSPKLLIYWASTR 60
Db 1 DIVISQSPSSLAWSAGEKVTMCKSSQSLNLSRTRKNYLAWYQQKPGQSPKLLIYWASTR 60
QY 61 ESGVPDRFSGSGGTDFLTITSSVQAEDLAVYYCHQYLSYTFGGGTTKLEIK 112
Db 61 ECGVPDRFTGSGGTDFLTITSSVQAEDLAVYYCKQSYNLYTFGGGTTKLEIK 112
RESULT 24
A53261
Ig kappa chain precursor V-J-C region - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 19-May-1994 #sequence_revision 19-May-1994 #text_change 21-Jan-2000
C;Accession: A53261
R;Cogne, M.; Preud'homme, J.L.; Bauwens, M.; Touchard, G.; Aucouturier, P.
J. Clin. Invest. 87, 2186-2190, 1991
A;Title: Structure of a monoclonal kappa chain of the Vk-IV subgroup in the kidney and p

A;Reference number: A53261; MUID:91250576; PMID:1904072
A;Accession: A53261
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-138 <COG>
A;Cross-references: GB:M38267
A;Note: authors translated the codon TTG for residue 12 as Phe
C;Superfamily: immunoglobulin V region; immunoglobulin homology
C;Keywords: heterotetramer; immunoglobulin
F;36-116/Domain: immunoglobulin homology <IMM>
Query Match 83.4%; Score 486; DB 2; Length 138;
Best Local Similarity 82.5%; Pred. No. 5.5e-37;
Matches 94; Conservative 8; Mismatches 10; Indels 2; Gaps 1;
QY 1 DIVMTQSPDLSAVSLGERVTMCKSSQSVLYSSNQKNYLAWYQQKPGQSPKLLIYWASTR 60
Db 21 DIVMTQSPDLSAVSLGERATINCKSSLSVFFSPNNKNYLAWYQQKPGQPPKLLIYWASTR 80
QY 61 ESGVPDRFSGSGGTDFLTITSSVQAEDLAVYYCHQYLS--SYTFGGGTTKLEIK 112
Db 81 ESGVPDRFSGSGGTDFLTITSLRQAEDVAVYYCQYTTLSWTFTGGGTTKVEIK 134
RESULT 25
S40364
Ig kappa chain - human
C;Species: Homo sapiens (man)
C;Date: 06-Mar-1994 #sequence_revision 26-May-1995 #text_change 21-Jan-2000
C;Accession: S40364
R;Klein, R.; Jaenichen, R.; Zachau, H.G.
Eur. J. Immunol. 23, 3248-3271, 1993
A;Title: Expressed human immunoglobulin chi genes and their hypermutation.
A;Reference number: S40312; MUID:94080891; PMID:8258341
A;Accession: S40364
A;Status: preliminary; translation not shown
A;Molecule type: mRNA
A;Residues: 1-124 <KLE>
A;Cross-references: EMBL:X72474; NID:G441416; PID:G441417
C;Superfamily: immunoglobulin V region; immunoglobulin homology
C;Keywords: heterotetramer; immunoglobulin
F;26-106/Domain: immunoglobulin homology <IMM>
Query Match 83.3%; Score 485.5; DB 2; Length 124;
Best Local Similarity 80.5%; Pred. No. 5.5e-37;
Matches 91; Conservative 12; Mismatches 9; Indels 1; Gaps 1;
QY 1 DIVMTQSPDLSAVSLGERVTMCKSSQSVLYSSNQKNYLAWYQQKPGQSPKLLIYWASTR 60
Db 11 DIVMTQSPDLSAVSLGERATINCKSSRSILYTSNNKNYLAWYQHKPGQPPRLLIYWASNR 70
QY 61 ESGVPDRFSGSGGTDFLTITSSVQAEDLAVYYCHQ-YLSSYTFGGGTTKLEIK 112
Db 71 ESGVPDRFSASGSGGTDFLTITSSVQAEDVAVYYCQYINPISFGGTTKVIK 123
Search completed: June 23, 2003, 14:04:14
Job time : 8.5461 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 23, 2003, 13:59:10 ; Search time 14.4965 Seconds
(without alignments)
1591.926 Million cell updates/sec

Title: US-10-056-052A-18
Perfect score: 583
Sequence: 1 DIVMTQSPDSLAVSLGERVT.....YCHQYLSSYTFGGGTKLEIK 112

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues
al number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

- Database : SPTREMBL_21.*
- 1: sp_archaea.*
 - 2: sp_bacteria.*
 - 3: sp_fungi.*
 - 4: sp_human.*
 - 5: sp_invertebrate.*
 - 6: sp_mammal.*
 - 7: sp_mhc.*
 - 8: sp_organelle.*
 - 9: sp_phage.*
 - 10: sp_plant.*
 - 11: sp_rodent.*
 - 12: sp_virus.*
 - 13: sp_vertebrate.*
 - 14: sp_unclassified.*
 - 15: sp_rvirus.*
 - 16: sp_bacteriap.*
 - 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	432.5	74.2	107	11	Q9ERZ9
2	391	67.1	109	4	Q9UL78
3	389	66.7	107	4	Q96SA9
4	383	65.7	238	11	Q99M37
5	378.5	64.9	108	4	Q9UL70
6	378	64.8	109	4	Q9UL85
7	377	64.7	238	11	Q8VC16
8	376.5	64.6	108	4	Q9UL77
9	376.5	64.6	111	11	Q920E9
10	376	64.5	107	4	Q9UL81
11	372.5	63.9	108	4	Q9UL83
12	371.5	63.7	108	4	Q9UL79
13	367	63.0	109	4	Q9UL86
14	366	62.8	239	11	Q8VC55
15	363.5	62.3	99	11	Q9JL74
16	363.5	62.3	108	11	Q8VIJ0

90 108 18.5 143 11 Q924R0 Q924r0 mus musculu
91 108 18.5 143 11 Q924P9 Q924p9 mus musculu
92 108 18.5 224 13 Q8UV32 Q8uv32 brachydanio
93 108 18.5 260 13 Q8UV96 Q8uv96 brachydanio
94 108 18.5 275 13 Q8UV91 Q8uv91 brachydanio
95 107.5 18.4 308 13 Q8UUG3 Q8uug3 ictalurus p
96 107 18.4 143 11 Q91V67 Q91v67 mus musculu
97 107 18.4 145 11 Q924R1 Q924r1 mus musculu
98 106.5 18.3 146 11 Q924Q3 Q924q3 mus musculu
99 106.5 18.3 168 4 Q9UQ56 Q9uq56 homo sapien
100 106.5 18.3 215 13 Q8UWK7 Q8uwk7 ictalurus p

ALIGNMENTS

RESULT 1
Q9ER29
ID Q9ER29 PRELIMINARY; PRT; 107 AA.
AC Q9ER29;
01-MAR-2001 (TrEMBLrel. 16, Created)
01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Anti human TNF-alpha light chain variable region (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Chen P., Deng J.B., Wang Z.L., Han H., Su C.Z.;
RT "Cloning and sequencing of the light chain fragment of variable region
RT genes of an anti-hTNF-a monoclonal antibody.";
RL J. Cell. Mol. Immunol. 12:21-26(1996).
RN [2]
RP SEQUENCE FROM N.A.
RA Chen P., Deng J.B., Wang Z.L., Han H., Yao L.B., Su C.Z.;
RT "Construction and sequencing of the single-chain antibody gene of a
RT human TNF-alpha specific monoclonal antibody.";
RL Ti 4 Chun i Ta Hsueh Hsueh Pao 19:373-376(1998).
RN [3]
RP SEQUENCE FROM N.A.
RA Chen P., Deng J.B., Wang Z.L., Han H., Yao L.B., Su C.Z.;
RL Submitted (MAY-2000) to the EMBL/GenBank/DBBJ databases.
DR EMBL; AF262753; AAG23804.1; -
DR HSSP; P80362; 1WTL.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
PFam; PF00047; ig; 1.
SMART; SM00409; IG; 1.
SMART; SM00406; IGV; 1.
FT NON_TER 107
SQ SEQUENCE 107 AA; 11784 MW; 2B15EEA6604A26C3 CRC64;

Query Match 74.2%; Score 432.5; DB 11; Length 107;
Best Local Similarity 77.6%; Pred. NO. 1.8e-38;
Matches 83; Conservative 13; Mismatches 10; Indels 1; Gaps 1;
QY 4 MTQSPDSLAVSLGERVTMCKSSQSVLYSSNQKNYLAQYQKPGQSPKLLIYWASTRESG 63
Db 1 MTQSPSSLAMSVGKVTMCKSSQSVLYSSNQKNYLAQYQKPGQSPKLLIYWASTRESG 60
QY 64 VPDRFSGSGGTDFLTITSSVQAEDLAVYVCHQ-YLSSYTFGGGTKL 109
Db 61 VPDRFMGSGSGGTDFLTITSSVQTEDLADYFCQHYRTPFTFGSGGTKL 107

RESULT 2
Q9UL78
ID Q9UL78 PRELIMINARY; PRT; 109 AA.
AC Q9UL78;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)

DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Myosin-reactive immunoglobulin light chain variable region
DE (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98277139; PubMed=9614934;
RA Wu X., Liu B., Van der Merwe P.L., Kalis N.N., Berney S.M.,
RA Young D.C.;
RT "Myosin-reactive autoantibodies in rheumatic carditis and normal
RT fetus.";
RL Clin. Immunol. Immunopathol. 87:184-192(1998).
DR EMBL; AF035036; AAD56272.1; -
DR HSSP; P80362; 1WTL.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; ig; 1.
DR SMART; SM00406; IGV; 1.
FT NON_TER 1
FT NON_TER 109
SQ SEQUENCE 109 AA; 11646 MW; 5F675C52EC7EE197 CRC64;
Query Match 67.1%; Score 391; DB 4; Length 109;
Best Local Similarity 67.3%; Pred. NO. 4.7e-34;
Matches 76; Conservative 18; Mismatches 13; Indels 6; Gaps 2;
QY 1 DIVMTQSPDSLAVSLGERVTMCKSSQSVLYSSNQKNYLAQYQKPGQSPKLLIYWASTR 60
Db 1 EIVLTQSPGTLSLSPGERATLSCRASQSV-----SSSYLAQYQKPGQAPRLLIYGASSR 55
QY 61 ESGVPDRFSGSGGTDFLTITSSVQAEDLAVYVCHQYLSS-YTFGGGTKLEIK 112
Db 56 ATGIPDRFSGSGGTDFLTITSLRLEPEDCAVYVCOQYGSPLTFGGGTKVEIK 108
RESULT 3
Q96SA9
ID Q96SA9 PRELIMINARY; PRT; 107 AA.
AC Q96SA9;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Anti-streptococcal/anti-myosin immunoglobulin kappa light chain
DE variable region (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98375893; PubMed=9712075;
RA Adderson E.E., Shikman A.R., Ward K.E., Cunningham M.W.;
RT "Molecular analysis of polyclonal monoclonal antibodies from
RT rheumatic carditis: human anti-N-acetylglucosamine/anti-myosin
RT antibody v region genes.";
RL J. Immunol. 161:2020-2031(1998).
DR EMBL; U96396; AAB68785.1; -
DR InterPro; IPR003006; Ig_MHC.
DR Pfam; PF00047; ig; 1.
FT NON_TER 1
FT NON_TER 107
SQ SEQUENCE 107 AA; 11520 MW; 4BB43E9C5B577F16 CRC64;
Query Match 66.7%; Score 389; DB 4; Length 107;
Best Local Similarity 67.0%; Pred. NO. 7.4e-34;
Matches 75; Conservative 15; Mismatches 16; Indels 6; Gaps 1;
QY 1 DIVMTQSPDSLAVSLGERVTMCKSSQSVLYSSNQKNYLAQYQKPGQSPKLLIYWASTR 60
Db 1 DIVMTQSPDSLAVSLGERVTMCKSSQSVLYSSNQKNYLAQYQKPGQSPKLLIYWASTR 60

Db 1 DIQMTQSPSSLSASVGDRTVTITCRASQSI-----SSYLNWYQQKPGKAPKLLIYAASSL 54

QY 61 ESGVPDRFSGSGGTDTLTITSSVQAEDLAVYYCHQYLSSTYTFGGGTKLEIK 112

Db 55 QSGVPSRFSGSGGTDTLTITSSLPEDFATYYCQSYSTLTTFGGGTKVEIK 106

RESULT 4

Q99M37

ID Q99M37 PRELIMINARY; PRT; 238 AA.

AC Q99M37;

DT 01-JUN-2001 (TrEMBLrel. 17, Created)

DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)

DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)

DE Hypothetical 26.3 kDa protein.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

[1]

SEQUENCE FROM N.A.

Strausberg R.;

Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.

RL EMBL; BC02035; AAH02035.1; -.

DR HSSP; P01679; 2FBJ.

DR InterPro; IPR003599; Ig.

DR InterPro; IPR003597; Ig_c1.

DR InterPro; IPR003600; Ig_like.

DR InterPro; IPR003006; Ig_MHC.

DR InterPro; IPR003596; Ig_v.

DR Pfam; PF00047; Ig; 2.

DR SMART; SM00409; IG; 2.

DR SMART; SM00407; IGc1; 1.

DR SMART; SM00406; IGv; 1.

DR SMART; SM00410; IG_like; 1.

DR PROSITE; PS00290; IG_MHC; UNKNOWN_1.

KW Hypothetical protein.

SQ SEQUENCE 238 AA; 26344 MW; FB2B06A0B801330A CRC64;

Query Match 65.7%; Score 383; DB 11; Length 238;

Best Local Similarity 69.0%; Pred. No. 9e-33;

Matches 78; Conservative 12; Mismatches 21; Indels 2; Gaps 2;

QY 1 DIVMTQSPDSLAVSLGERVTMNCSSQSVLYSSNQKNYLAWYQQKPGQSPKLLIYWASTR 60

Db 20 DVVMTQTPLSLPVSLGDAQISICRSSQSIH-SNGNTYLEWYLQKPGQSPKLLIYKVSNR 78

61 ESGVPDRFSGSGGTDTLTITSSVQAEDLAVYYCHQYLS-SYTFGGGTKLEIK 112

79 FSGVPDRFSGSGGTDTLTIKISRVEAEDLGVIYCFQGSHPYTFGGGTKLEIK 131

RESULT 5

Q9UL70

ID Q9UL70 PRELIMINARY; PRT; 108 AA.

AC Q9UL70;

DT 01-MAY-2000 (TrEMBLrel. 13, Created)

DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)

DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)

DE Myosin-reactive immunoglobulin light chain variable region

DE (Fragment).

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.

OX NCBI_TaxID=9606;

[1]

SEQUENCE FROM N.A.

MEDLINE=98277139; PubMed=9614934;

RA Wu X., Liu B., Van der Merwe P.L., Kalis N.N., Berney S.M.,

RA Young D.C.;

RT "Myosin-reactive autoantibodies in rheumatic carditis and normal fetus."

RL Clin. Immunol. Immunopathol. 87:184-192(1998).

DR EMBL; AF035044; AAD56280.1; -.

DR HSSP; P01607; 1REI.

DR InterPro; IPR003006; Ig_MHC.

DR InterPro; IPR003596; Ig_v.

DR Pfam; PF00047; Ig; 1.

DR SMART; SM00406; IGv; 1.

FT NON_TER 1

FT NON_TER 108

SQ SEQUENCE 108 AA; 11633 MW; B7BEDC3E41FCCA37 CRC64;

Query Match 64.9%; Score 378.5; DB 4; Length 108;

Best Local Similarity 67.3%; Pred. No. 9.8e-33;

Matches 76; Conservative 13; Mismatches 17; Indels 7; Gaps 2;

QY 1 DIVMTQSPDSLAVSLGERVTMNCSSQSVLYSSNQKNYLAWYQQKPGQSPKLLIYWASTR 60

Db 1 DIQMTQSPSSLSASVGDRTVTITCRASQGI-----SNYLAWYQQKPGKVPKSLIYAASL 54

QY 61 ESGVPDRFSGSGGTDTLTITSSVQAEDLAVYYCHQYLS-SYTFGGGTKLEIK 112

Db 55 QSGVPSRFSGSGGTDTLTITSSLPEDVATYYCQYNSAPRTFGGTKLEIK 107

RESULT 6

Q9UL85

ID Q9UL85 PRELIMINARY; PRT; 109 AA.

AC Q9UL85;

DT 01-MAY-2000 (TrEMBLrel. 13, Created)

DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)

DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)

DE Myosin-reactive immunoglobulin kappa chain variable region

DE (Fragment).

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=98277139; PubMed=9614934;

RA Wu X., Liu B., Van der Merwe P.L., Kalis N.N., Berney S.M.,

RA Young D.C.;

RT "Myosin-reactive autoantibodies in rheumatic carditis and normal fetus."

RL Clin. Immunol. Immunopathol. 87:184-192(1998).

DR EMBL; AF035029; AAD56265.1; -.

DR HSSP; P80362; 1WTL.

DR InterPro; IPR003006; Ig_MHC.

DR InterPro; IPR003596; Ig_v.

DR Pfam; PF00047; ig; 1.

DR SMART; SM00406; IGv; 1.

FT NON_TER 1

FT NON_TER 109

SQ SEQUENCE 109 AA; 11761 MW; FB1E43E7C7AFACCC CRC64;

Query Match 64.8%; Score 378; DB 4; Length 109;

Best Local Similarity 67.5%; Pred. No. 1.1e-32;

Matches 77; Conservative 17; Mismatches 12; Indels 8; Gaps 3;

QY 1 DIVMTQSPDSLAVSLGERVTMNCSSQSVLYSSNQKNYLAWYQQKPGQSPKLLIYWASTR 60

Db 1 EIVMTQSPATLSVSPGERATLSCWASQSI--SSN----LAWYQQKPGQAPRLIYGASTR 54

QY 61 ESGVPDRFSGSGGTDTLTITSSVQAEDLAVYYCHQYLS--SYTFGGGTKLEIK 112

Db 55 ATGIPARFSGSGGTDTLTITSSLPQEDFAIYHCQYNSWPPLTFGGGTKVEIK 108

RESULT 7

Q8VCI6

ID Q8VCI6 PRELIMINARY; PRT; 238 AA.

AC Q8VCI6;

DT 01-MAR-2002 (TrEMBLrel. 20, Created)

DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)

DR SMART; SM00406; IGV; 1.
FT NON_TER 1
FT NON_TER 107
SQ SEQUENCE 107 AA; 11501 MW; 070549FDE0754748 CRC64;

Query Match 64.5%; Score 376; DB 4; Length 107;
Best Local Similarity 64.3%; Pred. No. 1.8e-32;
Matches 72; Conservative 16; Mismatches 18; Indels 6; Gaps 1;

QY 1 DIVMTQSPDLSAVSLGERVTMCKSSQSVLYSSNQKNYLAWYQKPGQSPKLLIYWASTR 60
Db 1 DIQMTQSPSSLSASVGDRTITCRASQSI-----SNYLNWYQKPGKAPNLLIYAASSL 54

QY 61 ESGVPDRFSGSGGTDFTLTISSVQAEDLAVYVCHQYLSYTFGGGKLEIK 112
Db 55 QSGVPSRFSGSGGTDFTLTIISGLQAEDFATYVCOQSYALTGPGTKVDIR 106

RESULT 11
Q9UL83 PRELIMINARY; PRT; 108 AA.
Q9UL83;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Myosin-reactive immunoglobulin light chain variable region
DE (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98277139; PubMed=9614934;
RA Wu X., Liu B., Van der Merwe P.L., Kalis N.N., Berney S.M.,
RA Young D.C.;
RT "Myosin-reactive autoantibodies in rheumatic carditis and normal fetus."
RL Clin. Immunol. Immunopathol. 87:184-192(1998).
DR EMBL; AF035031; AAD56267.1; --
DR HSSP; P80362; 1WTL.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IGV; 1.
FT NON_TER 1
FT NON_TER 108
SQ SEQUENCE 108 AA; 11834 MW; 9F9C5A92EBA96EEA CRC64;

Query Match 63.9%; Score 372.5; DB 4; Length 108;
Best Local Similarity 67.3%; Pred. No. 4.3e-32;
Matches 76; Conservative 17; Mismatches 13; Indels 7; Gaps 3;

QY 1 DIVMTQSPDLSAVSLGERVTMCKSSQSVLYSSNQKNYLAWYQKPGQSPKLLIYWASTR 60
Db 1 EIVMTQSPATLSVSPGERATLSCRASQSV--SSN----LAWYQKPGQAPRLLIYCASTR 54

QY 61 ESGVPDRFSGSGGTDFTLTISSVQAEDLAVYVCHQYLS-SYTFGGGKLEIK 112
Db 55 ATGIPARFSGSGGTDFTLTIISLQFEDFAVYVCHQYNNWPFTEGPGTKVDIK 107

RESULT 12
Q9UL79 PRELIMINARY; PRT; 108 AA.
Q9UL79;
AC Q9UL79;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Myosin-reactive immunoglobulin light chain variable region
DE (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98277139; PubMed=9614934;
RA Wu X., Liu B., Van der Merwe P.L., Kalis N.N., Berney S.M.,
RA Young D.C.;
RT "Myosin-reactive autoantibodies in rheumatic carditis and normal fetus."
RL Clin. Immunol. Immunopathol. 87:184-192(1998).
DR EMBL; AF035035; AAD56271.1; --
DR HSSP; P01607; 1REI.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; ig; 1.
DR SMART; SM00406; IGV; 1.
FT NON_TER 1
FT NON_TER 108
SQ SEQUENCE 108 AA; 11787 MW; DB5845F19724FB4E CRC64;

Query Match 63.7%; Score 371.5; DB 4; Length 108;
Best Local Similarity 65.5%; Pred. No. 5.5e-32;
Matches 74; Conservative 14; Mismatches 18; Indels 7; Gaps 2;

QY 1 DIVMTQSPDLSAVSLGERVTMCKSSQSVLYSSNQKNYLAWYQKPGQSPKLLIYWASTR 60
Db 1 DIVMTQSPSLLSASTGDRVTISCRMSQGI-----SSYLAWYQKPGKAPELLIYAASL 54

QY 61 ESGVPDRFSGSGGTDFTLTISSVQAEDLAVYVCHQYLS-SYTFGGGKLEIK 112
Db 55 QSGVPSRFSGSGGTDFTLTIISCLQSEDFATYVCOQYYSFPPTFGGQTKVEIK 107

RESULT 13
Q9UL86 PRELIMINARY; PRT; 109 AA.
Q9UL86;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Myosin-reactive immunoglobulin kappa chain variable region
DE (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98277139; PubMed=9614934;
RA Wu X., Liu B., Van der Merwe P.L., Kalis N.N., Berney S.M.,
RA Young D.C.;
RT "Myosin-reactive autoantibodies in rheumatic carditis and normal fetus."
RL Clin. Immunol. Immunopathol. 87:184-192(1998).
DR EMBL; AF035028; AAD56264.1; --
DR HSSP; P80362; 1WTL.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; ig; 1.
DR SMART; SM00406; IGV; 1.
FT NON_TER 1
FT NON_TER 109
SQ SEQUENCE 109 AA; 11928 MW; 243325F72C7DAC83 CRC64;

Query Match 63.0%; Score 367; DB 4; Length 109;
Best Local Similarity 62.8%; Pred. No. 1.7e-31;
Matches 71; Conservative 20; Mismatches 16; Indels 6; Gaps 2;

QY 1 DIVMTQSPDLSAVSLGERVTMCKSSQSVLYSSNQKNYLAWYQKPGQSPKLLIYWASTR 60
Db 1 EIVLTQSPGTLTLSPGERATLSCRASQSV-----SSSYLAWYQKPGQAPRLLIYGTSSR 55

QY 61 ESGVPDRFSGSGGTDFTLTISSVQAEDLAVYVCHQYLS-SYTFGGGKLEIK 112

RL Submitted (SEP-2000) to the EMBL/GenBank/DBBJ databases.

DR EMBL; AF307938; AAL09422.1; -

DR InterPro; IPR003006; Ig_MHC.

DR Pfam; PF00047; Ig; 1.

FT NON_TER 1

FT NON_TER 109 109

SQ SEQUENCE 109 AA; 11943 MW; DAD3F98E05DDI501 CRC64;

Query Match 57.2%; Score 333.5; DB 11; Length 109;

Best Local Similarity 59.3%; Pred. No. 6e-28;

Matches 67; Conservative 15; Mismatches 24; Indels 7; Gaps 2;

QY 1 DIVMTQSPDLSAVSLGERVTMCKSSQSVLYSSNQKNYLAWYQQKPGQSPKLLIYWASTR 60
|| ||||| ||: ||||| ||: ||: || ||||| ||: ||: ||: ||: ||

Db 1 DIQMTQSPASLSASVGETVTTCRA-----SGNIHNYLAWYQQKQKSPQLLVYNAKTL 54
|| ||||| ||: ||||| ||: ||: || ||||| ||: ||: ||: ||: ||

QY 61 ESGVPDRFSGSGGTDFTLTISSVQAEDLAVYYCHQYLSS-YTFGGGTKLEIK 112
||| ||||| ||: ||: ||: || || || ||: ||: ||: ||: ||

Db 55 ADGVPDRFSGSGGTQYSLKINSIQPEDEFGSYQCQHFNSTPWTFTGGGTKLEIK 107
||| ||||| ||: ||: ||: || || || ||: ||: ||: ||: ||

QY 55 ADGVPDRFSGSGGTQYSLKINSIQPEDEFGSYQCQHFNSTPWTFTGGGTKLEIK 107

Db 55 ADGVPDRFSGSGGTQYSLKINSIQPEDEFGSYQCQHFNSTPWTFTGGGTKLEIK 107

QY 55 ADGVPDRFSGSGGTQYSLKINSIQPEDEFGSYQCQHFNSTPWTFTGGGTKLEIK 107

Db 55 ADGVPDRFSGSGGTQYSLKINSIQPEDEFGSYQCQHFNSTPWTFTGGGTKLEIK 107

QY 55 ADGVPDRFSGSGGTQYSLKINSIQPEDEFGSYQCQHFNSTPWTFTGGGTKLEIK 107

Db 55 ADGVPDRFSGSGGTQYSLKINSIQPEDEFGSYQCQHFNSTPWTFTGGGTKLEIK 107

QY 55 ADGVPDRFSGSGGTQYSLKINSIQPEDEFGSYQCQHFNSTPWTFTGGGTKLEIK 107

Db 55 ADGVPDRFSGSGGTQYSLKINSIQPEDEFGSYQCQHFNSTPWTFTGGGTKLEIK 107

QY 55 ADGVPDRFSGSGGTQYSLKINSIQPEDEFGSYQCQHFNSTPWTFTGGGTKLEIK 107

Db 55 ADGVPDRFSGSGGTQYSLKINSIQPEDEFGSYQCQHFNSTPWTFTGGGTKLEIK 107

QY 55 ADGVPDRFSGSGGTQYSLKINSIQPEDEFGSYQCQHFNSTPWTFTGGGTKLEIK 107

Db 55 ADGVPDRFSGSGGTQYSLKINSIQPEDEFGSYQCQHFNSTPWTFTGGGTKLEIK 107

QY 55 ADGVPDRFSGSGGTQYSLKINSIQPEDEFGSYQCQHFNSTPWTFTGGGTKLEIK 107

Db 55 ADGVPDRFSGSGGTQYSLKINSIQPEDEFGSYQCQHFNSTPWTFTGGGTKLEIK 107

QY 55 ADGVPDRFSGSGGTQYSLKINSIQPEDEFGSYQCQHFNSTPWTFTGGGTKLEIK 107

Db 55 ADGVPDRFSGSGGTQYSLKINSIQPEDEFGSYQCQHFNSTPWTFTGGGTKLEIK 107

QY 55 ADGVPDRFSGSGGTQYSLKINSIQPEDEFGSYQCQHFNSTPWTFTGGGTKLEIK 107

Db 55 ADGVPDRFSGSGGTQYSLKINSIQPEDEFGSYQCQHFNSTPWTFTGGGTKLEIK 107

QY 55 ADGVPDRFSGSGGTQYSLKINSIQPEDEFGSYQCQHFNSTPWTFTGGGTKLEIK 107

Db 55 ADGVPDRFSGSGGTQYSLKINSIQPEDEFGSYQCQHFNSTPWTFTGGGTKLEIK 107

QY 55 ADGVPDRFSGSGGTQYSLKINSIQPEDEFGSYQCQHFNSTPWTFTGGGTKLEIK 107

Db 55 ADGVPDRFSGSGGTQYSLKINSIQPEDEFGSYQCQHFNSTPWTFTGGGTKLEIK 107

QY 55 ADGVPDRFSGSGGTQYSLKINSIQPEDEFGSYQCQHFNSTPWTFTGGGTKLEIK 107

Db 55 ADGVPDRFSGSGGTQYSLKINSIQPEDEFGSYQCQHFNSTPWTFTGGGTKLEIK 107

QY 55 ADGVPDRFSGSGGTQYSLKINSIQPEDEFGSYQCQHFNSTPWTFTGGGTKLEIK 107

Db 55 ADGVPDRFSGSGGTQYSLKINSIQPEDEFGSYQCQHFNSTPWTFTGGGTKLEIK 107

QY 55 ADGVPDRFSGSGGTQYSLKINSIQPEDEFGSYQCQHFNSTPWTFTGGGTKLEIK 107

Db 55 ADGVPDRFSGSGGTQYSLKINSIQPEDEFGSYQCQHFNSTPWTFTGGGTKLEIK 107

QY 55 ADGVPDRFSGSGGTQYSLKINSIQPEDEFGSYQCQHFNSTPWTFTGGGTKLEIK 107

Db 55 ADGVPDRFSGSGGTQYSLKINSIQPEDEFGSYQCQHFNSTPWTFTGGGTKLEIK 107

QY 55 ADGVPDRFSGSGGTQYSLKINSIQPEDEFGSYQCQHFNSTPWTFTGGGTKLEIK 107

Db 55 ADGVPDRFSGSGGTQYSLKINSIQPEDEFGSYQCQHFNSTPWTFTGGGTKLEIK 107

QY 55 ADGVPDRFSGSGGTQYSLKINSIQPEDEFGSYQCQHFNSTPWTFTGGGTKLEIK 107

Db 55 ADGVPDRFSGSGGTQYSLKINSIQPEDEFGSYQCQHFNSTPWTFTGGGTKLEIK 107

QY 55 ADGVPDRFSGSGGTQYSLKINSIQPEDEFGSYQCQHFNSTPWTFTGGGTKLEIK 107

Db 55 ADGVPDRFSGSGGTQYSLKINSIQPEDEFGSYQCQHFNSTPWTFTGGGTKLEIK 107

QY 55 ADGVPDRFSGSGGTQYSLKINSIQPEDEFGSYQCQHFNSTPWTFTGGGTKLEIK 107

Db 55 ADGVPDRFSGSGGTQYSLKINSIQPEDEFGSYQCQHFNSTPWTFTGGGTKLEIK 107

QY 55 ADGVPDRFSGSGGTQYSLKINSIQPEDEFGSYQCQHFNSTPWTFTGGGTKLEIK 107

Db 55 ADGVPDRFSGSGGTQYSLKINSIQPEDEFGSYQCQHFNSTPWTFTGGGTKLEIK 107

QY 55 ADGVPDRFSGSGGTQYSLKINSIQPEDEFGSYQCQHFNSTPWTFTGGGTKLEIK 107

Db 55 ADGVPDRFSGSGGTQYSLKINSIQPEDEFGSYQCQHFNSTPWTFTGGGTKLEIK 107

QY 55 ADGVPDRFSGSGGTQYSLKINSIQPEDEFGSYQCQHFNSTPWTFTGGGTKLEIK 107

Db 55 ADGVPDRFSGSGGTQYSLKINSIQPEDEFGSYQCQHFNSTPWTFTGGGTKLEIK 107

QY 55 ADGVPDRFSGSGGTQYSLKINSIQPEDEFGSYQCQHFNSTPWTFTGGGTKLEIK 107

Db 55 ADGVPDRFSGSGGTQYSLKINSIQPEDEFGSYQCQHFNSTPWTFTGGGTKLEIK 107

QY 55 ADGVPDRFSGSGGTQYSLKINSIQPEDEFGSYQCQHFNSTPWTFTGGGTKLEIK 107

Db 55 ADGVPDRFSGSGGTQYSLKINSIQPEDEFGSYQCQHFNSTPWTFTGGGTKLEIK 107

QY 55 ADGVPDRFSGSGGTQYSLKINSIQPEDEFGSYQCQHFNSTPWTFTGGGTKLEIK 107

Db 55 ADGVPDRFSGSGGTQYSLKINSIQPEDEFGSYQCQHFNSTPWTFTGGGTKLEIK 107

QY 55 ADGVPDRFSGSGGTQYSLKINSIQPEDEFGSYQCQHFNSTPWTFTGGGTKLEIK 107

Db 55 ADGVPDRFSGSGGTQYSLKINSIQPEDEFGSYQCQHFNSTPWTFTGGGTKLEIK 107

Search completed: June 23, 2003, 14:03:30
Job time : 15.4965 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 23, 2003, 13:59:02 ; Search time 4.17021 Seconds
(without alignments)
1113.936 Million cell updates/sec

Title: US-10-056-052A-18
Perfect score: 583
Sequence: 1 DIVMTQSPDSLAVSLGERVT.....YCHQYLSSYTFGGGTKLEIK 112

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Minimum number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	521	89.4	133	1 KV4B_HUMAN	P06313 homo sapien
2	520.5	89.3	114	1 KV4A_HUMAN	P01625 homo sapien
3	505.5	86.7	134	1 KV4C_HUMAN	P06314 homo sapien
4	479	82.2	121	1 KV40_HUMAN	P06312 homo sapien
5	388.5	66.6	136	1 KV5B_MOUSE	P01634 mus musculu
6	387	66.4	110	1 KV3P_MOUSE	P01668 mus musculu
7	384.5	66.0	111	1 KV3L_MOUSE	P01664 mus musculu
8	381	65.4	129	1 KV3L_HUMAN	P18135 homo sapien
9	381	65.4	129	1 KV3M_HUMAN	P18136 homo sapien
10	379.5	65.1	149	1 KV5A_MOUSE	P01633 mus musculu
11	377	64.7	109	1 KV3B_HUMAN	P01620 homo sapien
12	376	64.5	109	1 KV3D_HUMAN	P01632 mus musculu
13	375.5	64.4	114	1 KV1A_MOUSE	P01622 homo sapien
14	374.5	64.2	108	1 KV1M_HUMAN	P01605 homo sapien
15	374	64.2	109	1 KV3E_HUMAN	P01623 homo sapien
16	373.5	64.1	111	1 KV3M_MOUSE	P01665 mus musculu
17	373.5	64.1	111	1 KV3O_MOUSE	P01667 mus musculu
18	372.5	63.9	111	1 KV3H_MOUSE	P01660 mus musculu
19	372	63.8	129	1 KV3H_HUMAN	P04207 homo sapien
20	367.5	63.0	111	1 KV3N_MOUSE	P01666 mus musculu
21	367.5	63.0	111	1 KV3R_MOUSE	P01670 mus musculu
22	367.5	63.0	129	1 KV1W_HUMAN	P04431 homo sapien
23	367	63.0	108	1 KV3A_HUMAN	P01619 homo sapien
24	367	63.0	109	1 KV3F_HUMAN	P01624 homo sapien
25	366.5	62.9	111	1 KV3J_MOUSE	P01662 mus musculu
26	366	62.8	133	1 KV2F_HUMAN	P06310 homo sapien
27	365.5	62.7	111	1 KV3Q_MOUSE	P01669 mus musculu
28	365.5	62.7	111	1 KV3U_MOUSE	P01673 mus musculu
29	365	62.6	113	1 KV2B_HUMAN	P01615 homo sapien
30	365	62.6	113	1 KV2G_MOUSE	P01631 mus musculu
31	364.5	62.5	108	1 KV1V_HUMAN	P04430 homo sapien
32	364.5	62.5	108	1 KV1Y_HUMAN	P80362 homo sapien
33	364.5	62.5	115	1 KV2A_HUMAN	P01614 homo sapien

ALIGNMENTS

RESULT 1

34	364	62.4	117	1	KV2E_HUMAN	P06309 homo sapien
35	362.5	62.2	108	1	KV1H_HUMAN	P01600 homo sapien
36	362	62.1	107	1	KV1D_HUMAN	P01596 homo sapien
37	362	62.1	109	1	KV3G_HUMAN	P04206 homo sapien
38	361.5	62.0	131	1	KV3I_MOUSE	P01661 mus musculu
39	360.5	61.8	108	1	KV1K_HUMAN	P01603 homo sapien
40	360.5	61.8	108	1	KV1Q_HUMAN	P01607 homo sapien
41	360.5	61.8	111	1	KV3S_MOUSE	P01671 mus musculu
42	360.5	61.8	128	1	KV3K_HUMAN	P06311 homo sapien
43	359	61.6	113	1	KV2D_HUMAN	P01617 homo sapien
44	358.5	61.5	111	1	KV3T_MOUSE	P01672 mus musculu
45	356.5	61.1	108	1	KV1F_HUMAN	P01598 homo sapien
46	353.5	60.6	108	1	KV1R_HUMAN	P01610 homo sapien
47	352.5	60.5	108	1	KV1B_HUMAN	P01594 homo sapien
48	351.5	60.3	108	1	KV1L_HUMAN	P01604 homo sapien
49	350.5	60.1	111	1	KV3D_MOUSE	P03977 mus musculu
50	350.5	60.1	111	1	KV3K_MOUSE	P01663 mus musculu
51	350.5	60.1	112	1	KV3G_MOUSE	P01659 mus musculu
52	349.5	59.9	108	1	KV1E_HUMAN	P01597 homo sapien
53	347.5	59.6	108	1	KV1P_HUMAN	P01608 homo sapien
54	347.5	59.6	112	1	KV2C_HUMAN	P01616 homo sapien
55	345	59.2	112	1	KV3B_MOUSE	P01655 mus musculu
56	344.5	59.1	108	1	KV5M_MOUSE	P01646 mus musculu
57	344	59.0	113	1	KV2E_MOUSE	P03976 mus musculu
58	343.5	58.9	108	1	KV1S_HUMAN	P01611 homo sapien
59	343.5	58.9	108	1	KV5P_MOUSE	P01649 mus musculu
60	343.5	58.9	132	1	KV3F_MOUSE	P01658 mus musculu
61	342.5	58.7	108	1	KV5J_MOUSE	P01643 mus musculu
62	341.5	58.6	108	1	KV1C_HUMAN	P01595 homo sapien
63	341.5	58.6	108	1	KV1Q_HUMAN	P01609 homo sapien
64	341.5	58.6	111	1	KV3C_MOUSE	P01656 mus musculu
65	341.5	58.6	111	1	KV3E_MOUSE	P01657 mus musculu
66	341.5	58.6	129	1	KV1X_HUMAN	P04432 homo sapien
67	340.5	58.4	108	1	KV1N_HUMAN	P01606 homo sapien
68	340.5	58.4	108	1	KV5S_MOUSE	P01652 mus musculu
69	340	58.3	113	1	KV2F_MOUSE	P01630 mus musculu
70	339.5	58.2	108	1	KV1G_HUMAN	P01599 homo sapien
71	339.5	58.2	108	1	KV3V_MOUSE	P01674 mus musculu
72	339.5	58.2	108	1	KV5Q_MOUSE	P01650 mus musculu
73	339.5	58.2	116	1	KV3J_HUMAN	P04434 homo sapien
74	339.5	58.2	128	1	KV5E_MOUSE	P01637 mus musculu
75	338.5	58.1	111	1	KV3A_MOUSE	P01654 mus musculu
76	337.5	57.9	100	1	KV3C_HUMAN	P01621 homo sapien
77	337.5	57.9	108	1	KV1A_HUMAN	P01593 homo sapien
78	337.5	57.9	108	1	KV5T_MOUSE	P01653 mus musculu
79	336.5	57.7	112	1	KV1U_HUMAN	P01613 homo sapien
80	336.5	57.7	117	1	KV1J_HUMAN	P01602 homo sapien
81	335	57.5	112	1	KV2D_MOUSE	P01629 mus musculu
82	333.5	57.2	108	1	KV5D_MOUSE	P01636 mus musculu
83	332.5	57.0	108	1	KV5K_MOUSE	P01644 mus musculu
84	332.5	57.0	108	1	KV5L_MOUSE	P01645 mus musculu
85	332.5	57.0	110	1	KV15_RABIT	P01696 oryctolagus
86	331	56.8	115	1	KV3I_HUMAN	P04433 homo sapien
87	330.5	56.7	108	1	KV5N_MOUSE	P01647 mus musculu
88	330.5	56.7	108	1	KV5O_MOUSE	P01648 mus musculu
89	330	56.6	109	1	KV1T_HUMAN	P01612 homo sapien
90	329	56.4	113	1	KV2C_MOUSE	P01628 mus musculu
91	327.5	56.2	108	1	KV5R_MOUSE	P01651 mus musculu
92	325	55.7	129	1	KV4A_MOUSE	P01680 mus musculu
93	322.5	55.3	108	1	KV06_RABIT	P01687 oryctolagus
94	322	55.2	112	1	KV2A_MOUSE	P01626 mus musculu
95	321.5	55.1	109	1	KV03_RABIT	P01684 oryctolagus
96	321	55.1	114	1	KV16_RABIT	P01697 oryctolagus
97	320.5	55.0	117	1	KV1I_HUMAN	P01601 homo sapien
98	320	54.9	107	1	KV6A_MOUSE	P01675 mus musculu
99	318.5	54.6	108	1	KV1_CANFA	P01618 canis famil
100	318	54.5	107	1	KV6D_MOUSE	P01678 mus musculu

```
KV4B_HUMAN
ID   KV4B_HUMAN          STANDARD;          PRT;          133 AA.
AC   P06313;
DT   01-JAN-1988 (Rel. 06, Created)
DT   01-JAN-1988 (Rel. 06, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   Ig kappa chain V-IV region JI precursor.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=86041853; PubMed=2997712;
RA   Klobeck H.G., Bornkamm G.W., Combriato G., Mocikat R., Pohlenz H.D.,
RA   Zachau H.G.;
RT   "Subgroup IV of human immunoglobulin K light chains is encoded by a
RT   single germline gene."
RL   Nucleic Acids Res. 13:6515-6529(1985).
CC   -----
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----
DR   EMBL; Z00022; CAA77317.1; -.
DR   PIR; A01904; K4HUJ1.
DR   HSSP; P80362; 1WFL.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR003596; Ig_V.
DR   Pfam; PF00047; Ig; 1.
DR   SMART; SM00406; IGV; 1.
KW   Immunoglobulin V region; Signal.
FT   SIGNAL              1      20
FT   CHAIN               21     133   IG KAPPA CHAIN V-IV REGION JI.
FT   DOMAIN             21     43     FRAMEWORK-1.
FT   DOMAIN             44     60     COMPLEMENTARITY-DETERMINING-1.
FT   DOMAIN             61     75     FRAMEWORK-2.
FT   DOMAIN             76     82     COMPLEMENTARITY-DETERMINING-2.
FT   DOMAIN             83     114    FRAMEWORK-3.
FT   DOMAIN             115    122    COMPLEMENTARITY-DETERMINING-3.
FT   DOMAIN             123    132    FRAMEWORK-4.
FT   DISULFID           43     114    BY SIMILARITY.
FT   NON_TER            133     133
SQ   SEQUENCE 133 AA; 14632 MW; 5FB3953066744AF4 CRC64;

Query Match      89.4%; Score 521; DB 1; Length 133;
Best Local Similarity 89.3%; Pred. No. 6.3e-49;
Matches 100; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY   1 DIVMTQSPDSLAVSLGERVTMCKSSQSVLYSSNQKNYLAWYQKPGQSPKLLIYWASTR 60
Db   21 DIVMTQSPDSLAVSLGERATINCKSSQSVLYSSNNKNYLAWYQKPGQPKLLIYWASTR 80
QY   61 ESGVPDRFSGSGGTDTFTLTISVQAEDLAVYYCHQYLSSYTFGGGTKLEIK 112
Db   81 ESGVPDRFSGSGGTDTFTLTISLQAEDVAVYYCQYDTIPTFGGGTKVEIK 132

RESULT 2
KV4A_HUMAN
ID   KV4A_HUMAN          STANDARD;          PRT;          114 AA.
AC   P01625;
DT   21-JUL-1986 (Rel. 01, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   Ig kappa chain V-IV region Len.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

KV4B_HUMAN
ID   KV4B_HUMAN          STANDARD;          PRT;          133 AA.
AC   P06313;
DT   01-JAN-1988 (Rel. 06, Created)
DT   01-JAN-1988 (Rel. 06, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   Ig kappa chain V-IV region JI precursor.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=76004342; PubMed=50995;
RA   Schneider M., Hilschmann N.;
RT   "The primary structure of a monoclonic immunoglobulin-L-chain of
RT   subgroup IV of the kappa type (Bence-Jones protein Len).";
RL   Hoppe-Seyler's Z. Physiol. Chem. 356:507-557(1975).
RN   [2]
RP   REVISION TO 9.
RA   Salomon A.;
RL   Submitted (AUG-1996) to the SWISS-PROT data bank.
CC   -!- MISCELLANEOUS: THE C REGION OF THIS CHAIN HAS THE INV (3) MARKER.
CC   -!- MISCELLANEOUS: THIS IS A BENCE-JONES PROTEIN.
DR   PIR; A01903; K4HULN.
DR   HSSP; P01607; IREI.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR003596; Ig_V.
DR   Pfam; PF00047; ig; 1.
DR   SMART; SM00406; IGV; 1.
KW   Immunoglobulin V region; Bence-Jones protein.
FT   DOMAIN             1      23     FRAMEWORK-1.
FT   DOMAIN             24     40     COMPLEMENTARITY-DETERMINING-1.
FT   DOMAIN             41     55     FRAMEWORK-2.
FT   DOMAIN             56     62     COMPLEMENTARITY-DETERMINING-2.
FT   DOMAIN             63     94     FRAMEWORK-3.
FT   DOMAIN             95     101    COMPLEMENTARITY-DETERMINING-3.
FT   DOMAIN             102    113    FRAMEWORK-4.
FT   DISULFID           23     94     BY SIMILARITY.
FT   NON_TER            114     114
SQ   SEQUENCE 114 AA; 12640 MW; 0647F1D17F236485 CRC64;

Query Match      89.3%; Score 520.5; DB 1; Length 114;
Best Local Similarity 89.4%; Pred. No. 5.9e-49;
Matches 101; Conservative 5; Mismatches 6; Indels 1; Gaps 1;

QY   1 DIVMTQSPDSLAVSLGERVTMCKSSQSVLYSSNQKNYLAWYQKPGQSPKLLIYWASTR 60
Db   1 DIVMTQSPDSLAVSLGERATINCKSSQSVLYSSNNKNYLAWYQKPGQPKLLIYWASTR 60
QY   61 ESGVPDRFSGSGGTDTFTLTISVQAEDLAVYYCHQYLSS-YTFGGGTKLEIK 112
Db   61 ESGVPDRFSGSGGTDTFTLTISLQAEDVAVYYCQYYSTPYSGGQGTKLEIK 113

RESULT 3
KV4C_HUMAN
ID   KV4C_HUMAN          STANDARD;          PRT;          134 AA.
AC   P06314;
DT   01-JAN-1988 (Rel. 06, Created)
DT   01-APR-1988 (Rel. 07, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   Ig kappa chain V-IV region B17 precursor.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=86041854; PubMed=2997713;
RA   Marsh P., Mills F., Gould H.;
RT   "Detection of a unique human V kappa IV germline gene by a cloned
RT   cDNA probe."
RL   Nucleic Acids Res. 13:6531-6544(1985).
RN   [2]
RP   REVISION TO 76.
RA   Marsh P.;
RL   Submitted (OCT-1986) to the EMBL/GenBank/DBJ databases.
CC   -----
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
```

```
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=76004342; PubMed=50995;
RA   Schneider M., Hilschmann N.;
RT   "The primary structure of a monoclonic immunoglobulin-L-chain of
RT   subgroup IV of the kappa type (Bence-Jones protein Len).";
RL   Hoppe-Seyler's Z. Physiol. Chem. 356:507-557(1975).
RN   [2]
RP   REVISION TO 9.
RA   Salomon A.;
RL   Submitted (AUG-1996) to the SWISS-PROT data bank.
CC   -!- MISCELLANEOUS: THE C REGION OF THIS CHAIN HAS THE INV (3) MARKER.
CC   -!- MISCELLANEOUS: THIS IS A BENCE-JONES PROTEIN.
DR   PIR; A01903; K4HULN.
DR   HSSP; P01607; IREI.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR003596; Ig_V.
DR   Pfam; PF00047; ig; 1.
DR   SMART; SM00406; IGV; 1.
KW   Immunoglobulin V region; Bence-Jones protein.
FT   DOMAIN             1      23     FRAMEWORK-1.
FT   DOMAIN             24     40     COMPLEMENTARITY-DETERMINING-1.
FT   DOMAIN             41     55     FRAMEWORK-2.
FT   DOMAIN             56     62     COMPLEMENTARITY-DETERMINING-2.
FT   DOMAIN             63     94     FRAMEWORK-3.
FT   DOMAIN             95     101    COMPLEMENTARITY-DETERMINING-3.
FT   DOMAIN             102    113    FRAMEWORK-4.
FT   DISULFID           23     94     BY SIMILARITY.
FT   NON_TER            114     114
SQ   SEQUENCE 114 AA; 12640 MW; 0647F1D17F236485 CRC64;

Query Match      89.3%; Score 520.5; DB 1; Length 114;
Best Local Similarity 89.4%; Pred. No. 5.9e-49;
Matches 101; Conservative 5; Mismatches 6; Indels 1; Gaps 1;

QY   1 DIVMTQSPDSLAVSLGERVTMCKSSQSVLYSSNQKNYLAWYQKPGQSPKLLIYWASTR 60
Db   1 DIVMTQSPDSLAVSLGERATINCKSSQSVLYSSNNKNYLAWYQKPGQPKLLIYWASTR 60
QY   61 ESGVPDRFSGSGGTDTFTLTISVQAEDLAVYYCHQYLSS-YTFGGGTKLEIK 112
Db   61 ESGVPDRFSGSGGTDTFTLTISLQAEDVAVYYCQYYSTPYSGGQGTKLEIK 113

RESULT 3
KV4C_HUMAN
ID   KV4C_HUMAN          STANDARD;          PRT;          134 AA.
AC   P06314;
DT   01-JAN-1988 (Rel. 06, Created)
DT   01-APR-1988 (Rel. 07, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   Ig kappa chain V-IV region B17 precursor.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=86041854; PubMed=2997713;
RA   Marsh P., Mills F., Gould H.;
RT   "Detection of a unique human V kappa IV germline gene by a cloned
RT   cDNA probe."
RL   Nucleic Acids Res. 13:6531-6544(1985).
RN   [2]
RP   REVISION TO 76.
RA   Marsh P.;
RL   Submitted (OCT-1986) to the EMBL/GenBank/DBJ databases.
CC   -----
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
```



```
CC LEUKEMIA.
DR PIR; PL0022; K3HUHA.
DR HSSP; P80362; 1WTL.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; Igv; 1.
KW Immunoglobulin V region; Signal.
FT SIGNAL 1 20
FT CHAIN 21 129 IG KAPPA CHAIN V-III REGION HAH.
FT DOMAIN 21 43 FRAMEWORK-1.
FT DOMAIN 44 55 COMPLEMENTARITY-DETERMINING-1.
FT DOMAIN 56 70 FRAMEWORK-2.
FT DOMAIN 71 77 COMPLEMENTARITY-DETERMINING-2.
FT DOMAIN 78 109 FRAMEWORK-3.
FT DOMAIN 110 118 COMPLEMENTARITY-DETERMINING-3.
FT DOMAIN 119 129 JK1 SEGMENT.
FT DISULFID 43 109 BY SIMILARITY.
FT NON_TER 129
SQ SEQUENCE 129 AA; 14073 MW; D3C55292772774D0 CRC64;

Query Match 65.4%; Score 381; DB 1; Length 129;
Best Local Similarity 65.5%; Pred. No. 6e-34;
Matches 74; Conservative 18; Mismatches 15; Indels 6; Gaps 2;

QY 1 DIVMTQSPDSLAVSLGERVTMCKSSQSVLYSSNQKNYLAQYQKPGQSPKLLIYWASTR 60
Db 21 EIVLTQSPGTLSPGERATLSCRASQSV-----SSSYLAQYQKPGQAPRLLIYGASSR 75

QY 61 ESGVPDRFSGSGSGTDFTLTISVQAEDLAVYYCHQY-LSSYTFGGGKLEIK 112
Db 76 ATGIPDRFSGSGSGTDFTLTISRLEPEDFAVYYCQYGTSPRTFGQGTKEIK 128

RESULT 9
KV3M_HUMAN STANDARD; PRT; 129 AA.
AC P18136;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig kappa chain V-III region HIC precursor.
OS Homo sapiens. (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
MEDLINE=88171307; PubMed=3127527;
Kipps T.J., Tomhave E., Chen P.P., Carson D.A.;
"Autoantibody-associated kappa light chain variable region gene
expressed in chronic lymphocytic leukemia with little or no somatic
mutation. Implications for etiology and immunotherapy.";
J. Exp. Med. 167:840-852(1988).
CC -I- DISEASE: THE PROTEIN IS ONE OF THE SURFACE IMMUNOGLOBULIN M
AUTOANTIBODIES EXPRESSED IN PATIENTS WITH CHRONIC LYMPHOCYTIC
LEUKEMIA.
DR PIR; PL0021; K3HUHI.
DR HSSP; P80362; 1WTL.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; Igv; 1.
KW Immunoglobulin V region; Signal.
FT SIGNAL 1 20
FT CHAIN 21 129 IG KAPPA CHAIN V-III REGION HIC.
FT DOMAIN 21 43 FRAMEWORK-1.
FT DOMAIN 44 55 COMPLEMENTARITY-DETERMINING-1.
FT DOMAIN 56 70 FRAMEWORK-2.
FT DOMAIN 71 77 COMPLEMENTARITY-DETERMINING-2.
FT DOMAIN 78 109 FRAMEWORK-3.
FT DOMAIN 110 118 COMPLEMENTARITY-DETERMINING-3.
FT DOMAIN 119 129 JK1 SEGMENT.
```

```
FT DISULFID 43 109 BY SIMILARITY.
FT NON_TER 129
SQ SEQUENCE 129 AA; 14070 MW; 7395528EA2BB74D6 CRC64;

Query Match 65.4%; Score 381; DB 1; Length 129;
Best Local Similarity 65.5%; Pred. No. 6e-34;
Matches 74; Conservative 19; Mismatches 14; Indels 6; Gaps 2;

QY 1 DIVMTQSPDSLAVSLGERVTMCKSSQSVLYSSNQKNYLAQYQKPGQSPKLLIYWASTR 60
Db 21 EIVLTQSPGTLSPGERATLSCRASQSV-----SSSYLAQYQKPGQAPRLLIYGASSR 75

QY 61 ESGVPDRFSGSGSGTDFTLTISVQAEDLAVYYCHQY-LSSYTFGGGKLEIK 112
Db 76 ATGIPDRFSGSGSGTDFTLTISRLEPEDFAVYYCQYGTSPRTFGQGTKEIK 128

RESULT 10
KV5A_MOUSE STANDARD; PRT; 149 AA.
ID KV5A_MOUSE
AC P01633;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig kappa chain V-V region MPC11 precursor.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE OF 1-71 FROM N.A.
MEDLINE=83001944; PubMed=6288267;
Kelley D.E., Coleclough C., Perry R.P.;
"Functional significance and evolutionary development of the
5'-terminal regions of immunoglobulin variable-region genes.";
Cell 29:681-689(1982).
RN [2]
RP SEQUENCE OF 41-149 FROM N.A.
MEDLINE=80176554; PubMed=6245773;
Rabbitts T.H., Hamlyn P.H., Matthysens G., Roe B.A.;
"The variability, arrangement, and rearrangement of immunoglobulin
genes.";
Can. J. Biochem. 58:176-187(1980).
RN [3]
RP SEQUENCE OF 30-149.
MEDLINE=78186617; PubMed=418775;
Smith G.P.;
"Sequence of the full-length immunoglobulin kappa-chain of mouse
myeloma MPC 11.";
Biochem. J. 171:337-347(1978).
CC -I- MISCELLANEOUS: THE MATURE CHAIN HAS 12 ADDITIONAL RESIDUES AT ITS
AMINO END, DUE TO A TANDEM DUPLICATION OF 36 NUCLEOTIDES AFTER THE
CODON FOR RESIDUE 36. RESIDUE 42 CORRESPONDS TO THE AMINO-TERMINAL
RESIDUE OF TYPICAL KAPPA CHAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; J00561; AAA38776.1; -.
DR PIR; A01916; KVMS11.
DR HSSP; P80362; 1WTL.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; Igv; 1.
KW Immunoglobulin V region; Signal; Repeat.
FT SIGNAL 1 29
FT CHAIN 30 149 IG KAPPA CHAIN V-V REGION MPC11.
```

```

FT DOMAIN 42 64 FRAMEWORK-1.
FT DOMAIN 65 75 COMPLEMENTARITY-DETERMINING-1.
FT DOMAIN 76 90 FRAMEWORK-2.
FT DOMAIN 91 97 COMPLEMENTARITY-DETERMINING-2.
FT DOMAIN 98 129 FRAMEWORK-3.
FT DOMAIN 130 138 COMPLEMENTARITY-DETERMINING-3.
FT DOMAIN 139 148 FRAMEWORK-4.
FT REPEAT 26 35
FT REPEAT 38 47
FT NON_TER 149
SQ SEQUENCE 149 AA; 16434 MW; B0480C87B682AC3E CRC64;

Query Match 65.1%; Score 379.5; DB 1; Length 149;
Best Local Similarity 69.0%; Pred. No. 1e-33;
Matches 78; Conservative 12; Mismatches 16; Indels 7; Gaps 2;

QY 1 DIVMTQSPDLSAVSLGERVTMCKSSQSVLYSSNQKNYLAHYQQKPGQSPKLLIYWASTR 60
Db 42 DIVMTQSHKFMSTSVGDRVSITCKASQDV-----STTVAWYQQKPGQSPKLLIYSASYR 95
61 ESGVPDRFSGSGSGTDFTLTISVQAEDLAVYCHQYLSS-YTFGGGTKLEIK 112
Db 96 YGVDPDRFTGSGSGTDFTLTISVQAEDLAVYCHQYHSTPPTFGGGTKLEIK 148

RESULT 11
KV3B_HUMAN STANDARD; PRT; 109 AA.
AC P01620;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig kappa chain V-III region SIE.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
SEQUENCE.
MEDLINE=82046598; PubMed=6794615;
Andrews D.W., Capra J.D.;
"Amino acid sequence of the variable regions of light chains from two
idiotypically cross-reactive human IgM anti-gamma-globulins of the Wa
group.";
Biochemistry 20:5816-5822(1981).
-!- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM AN IGM WITH ANTI-GAMMA
GLOBULIN ACTIVITY.
PIR; A01892; K3HUSI.
HSSP; P80362; 1WTL.
InterPro; IPR003006; Ig_MHC.
InterPro; IPR003596; Ig_V.
Pfam; PF00047; Ig; 1.
SMART; SM00406; IGV; 1.
Immunoglobulin V region.
BY SIMILARITY.
DISULFID 23 89
NON_TER 109
SEQUENCE 109 AA; 11775 MW; 7689C3ECD646FFB4 CRC64;

Query Match 64.7%; Score 377; DB 1; Length 109;
Best Local Similarity 64.6%; Pred. No. 1.3e-33;
Matches 73; Conservative 20; Mismatches 14; Indels 6; Gaps 2;

QY 1 DIVMTQSPDLSAVSLGERVTMCKSSQSVLYSSNQKNYLAHYQQKPGQSPKLLIYWASTR 60
Db 1 EIVLTQSPGTLSPGERATLSCRASQSV-----SNSYLAHYQQKPGQAPRLIYGASSR 55
61 ESGVPDRFSGSGSGTDFTLTISVQAEDLAVYCHQYLSS-YTFGGGTKLEIK 112
Db 56 ATGIPDRFSGSGSGTDFTLTISRLEPDFAVYCYQQYGGSSPQTFFGQSKVEIK 108

RESULT 12
KV3D_HUMAN
```

```

ID KV3D_HUMAN STANDARD; PRT; 109 AA.
AC P01622;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig kappa chain V-III region TI.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
SEQUENCE.
MEDLINE=72188439; PubMed=5027703;
Suter L., Barnikol H.U., Watanabe S., Hilschmann N.;
"Rule of antibody structure. The primary structure of a monoclonal
immunoglobulin L-chain of kappa-type, subgroup 3 (Bence-Jones protein
Ti). IV. The complete amino acid sequence and its significance for
the mechanism of antibody production.";
Hoppe-Seyler's Z. Physiol. Chem. 353:189-208(1972).
-!- MISCELLANEOUS: THE C REGION OF THIS CHAIN HAS THE INV (3) MARKER.
-!- MISCELLANEOUS: THIS IS A BENICE-JONES PROTEIN.
PIR; A01895; K3HUTI.
HSSP; P80362; 1WTL.
InterPro; IPR003006; Ig_MHC.
InterPro; IPR003596; Ig_V.
Pfam; PF00047; Ig; 1.
SMART; SM00406; IGV; 1.
Immunoglobulin V region; Bence-Jones protein.
BY SIMILARITY.
DISULFID 23 89
NON_TER 109
SEQUENCE 109 AA; 11788 MW; 8C35058CDC7749BC CRC64;

Query Match 64.5%; Score 376; DB 1; Length 109;
Best Local Similarity 64.6%; Pred. No. 1.7e-33;
Matches 73; Conservative 20; Mismatches 14; Indels 6; Gaps 2;

QY 1 DIVMTQSPDLSAVSLGERVTMCKSSQSVLYSSNQKNYLAHYQQKPGQSPKLLIYWASTR 60
Db 1 EIVLTQSPGTLSPGERATLSCRASQSV-----SNSFLAWYQQKPGQAPRLIYVASSR 55
61 ESGVPDRFSGSGSGTDFTLTISVQAEDLAVYCHQYLSS-YTFGGGTKLEIK 112
Db 56 ATGIPDRFSGSGSGTDFTLTISRLEPDFAVYCYQQYGGSSPSTFGQGTKVELK 108

RESULT 13
KVIA_MOUSE STANDARD; PRT; 114 AA.
AC P01632;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig kappa chain V-I region S107A.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
SEQUENCE FROM N.A.
MEDLINE=81241357; PubMed=6788890;
Kwan S.-P., Rudikoff S., Seidman J.G., Leder P., Scharff M.D.;
"Nucleic acid and protein sequences of phosphocholine-binding light
chains.";
J. Exp. Med. 153:1366-1370(1981).
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
CC
```

DR EMBL; U29423; AAC00033.1; -.
DR PIR; A01915; KVMS7A.
DR HSSP; P01607; IREI.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IGV; 1.
KW Immunoglobulin V region.
FT DOMAIN 1 23
FT DOMAIN 24 40
FT DOMAIN 41 55
FT DOMAIN 56 62
FT DOMAIN 63 94
FT DOMAIN 95 103
FT DOMAIN 104 113
FT DISULFID 23 94
FT NON_TER 114 114
SQ SEQUENCE 114 AA; 12717 MW; 32008EC8E9DBE67B CRC64;

Query Match 64.4%; Score 375.5; DB 1; Length 114;
Best Local Similarity 69.3%; Pred. No. 2e-33;
Matches 79; Conservative 12; Mismatches 20; Indels 3; Gaps 3;

QY 1 DIVMTQSPDSLAVSLGERVTMCKSSQSVLYSSNQK-NYLAWYQQKPGQSPKLLIYWAST 59
Db 1 DIVMTQSPFLAVTASKKVTISCTASES-LYSSKHVHYLAWYQKKPEQSPKLLIYGASN 59
QY 60 RESGVPDRFSGSGSGTDFTLTITSSVQAEDLAVYCHQYLS-SYTFGGGKLEIK 112
Db 60 RYIGVPDRFTGSGSGTDFTLTITSSVQVEDLTHYCAQFYSYPLTFGAGTKLELK 113

RESULT 14
KV1M_HUMAN STANDARD; PRT; 108 AA.
ID KV1M_HUMAN
AC P01605;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig kappa chain V-I region Lay.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE.
RX MEDLINE=77038198; PubMed=824717;
RA Capra J.D., Klapper D.G.;
RT "Complete amino acid sequence of the variable domains of two human
IgM anti-gamma globulins (Lay/Pom) with shared idiotypic
specificities.";
RL Scand. J. Immunol. 5:677-684(1976).
CC -!- MISCELLANEOUS: THE SECOND AND THIRD HYPERVARIABLE REGIONS OF THIS
CHAIN ARE IDENTICAL WITH THOSE OF THE HUMAN POM V-III KAPPA CHAIN,
WITH WHICH IT SHARES CERTAIN IDIOTYPIC DETERMINANTS.
CC -!- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM AN IGM WITH ANTI-GAMMA
GLOBULIN ACTIVITY.
DR PIR; A01871; K1HULY.
DR HSSP; P01607; IREI.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IGV; 1.
KW Immunoglobulin V region.
FT DOMAIN 1 23
FT DOMAIN 24 34
FT DOMAIN 35 49
FT DOMAIN 50 56
FT DOMAIN 57 88
FT DOMAIN 89 97
FT DOMAIN 98 107
FT DISULFID 23 88
FT NON_TER 108 108

FRAMEWORK-1.
COMPLEMENTARITY-DETERMINING-1.
FRAMEWORK-2.
COMPLEMENTARITY-DETERMINING-2.
FRAMEWORK-3.
COMPLEMENTARITY-DETERMINING-3.
FRAMEWORK-4.
BY SIMILARITY.

SQ SEQUENCE 108 AA; 11834 MW; 739993A95431434A CRC64;

Query Match 64.2%; Score 374.5; DB 1; Length 108;
Best Local Similarity 66.4%; Pred. No. 2.4e-33;
Matches 75; Conservative 14; Mismatches 17; Indels 7; Gaps 2;

QY 1 DIVMTQSPDSLAVSLGERVTMCKSSQSVLYSSNQK-NYLAWYQQKPGQSPKLLIYWASTR 60
Db 1 DIQMTQSPSSLVSVDRTITCQASQNV-----NAYLNWYQQKPGAPKLLIYGASTR 54
QY 61 ESGVPDRFSGSGSGTDFTLTITSSVQAEDLAVYCHQYLS-SYTFGGGKLEIK 112
Db 55 EAGVPSRFSGSGSGTDFTTFTISSLQPEDIAITYCQYNNWPPPTFGQGTKEVK 107

RESULT 15
KV3E_HUMAN STANDARD; PRT; 109 AA.
ID KV3E_HUMAN
AC P01623;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig kappa chain V-III region WOL.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE.
RX MEDLINE=82046598; PubMed=6794615;
RA Andrews D.W., Capra J.D.;
RT "Amino acid sequence of the variable regions of light chains from two
idiotypically cross-reactive human IgM anti-gamma-globulins of the Wa
group.";
RL Biochemistry 20:5816-5822(1981).
CC -!- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM AN IGM WITH ANTI-GAMMA
GLOBULIN ACTIVITY.
DR PIR; A01896; K3HUWL.
DR HSSP; P80362; IWTLL.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IGV; 1.
KW Immunoglobulin V region.
FT DISULFID 23 89
FT NON_TER 109 109
SQ SEQUENCE 109 AA; 11746 MW; 566C115E6B9CBEEE CRC64;

Query Match 64.2%; Score 374; DB 1; Length 109;
Best Local Similarity 64.6%; Pred. No. 2.8e-33;
Matches 73; Conservative 17; Mismatches 17; Indels 6; Gaps 2;

QY 1 DIVMTQSPDSLAVSLGERVTMCKSSQSVLYSSNQK-NYLAWYQQKPGQSPKLLIYWASTR 60
Db 1 EIVLTQSPGTLSPGERATLSCRASQSV-----SSGYLGWYQQKPGQAPRLIYGASSR 55
QY 61 ESGVPDRFSGSGSGTDFTLTITSSVQAEDLAVYCHQYLS-SYTFGGGKLEIK 112
Db 56 ATGIPDRFSGSGSGTDFTLTISRLEPEDFAVYCCQYGSGLRTFGQGTKEIK 108

RESULT 16
KV3M_MOUSE STANDARD; PRT; 111 AA.
ID KV3M_MOUSE
AC P01665;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig kappa chain V-III region PC 7043.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;


```
RN SEQUENCE.
RX MEDLINE=79073152; PubMed=103003;
RA Weigert M., Gatmaitan L., Loh E., Schilling J., Hood L.E.;
RT "Rearrangement of genetic information may produce immunoglobulin
diversity.";
RL Nature 276:785-790(1978).
DR PIR; A01937; KVM543.
DR HSSP; P80362; 1WTL.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IGV; 1.
KW Immunoglobulin V region.
FT DOMAIN 1 23
FT DOMAIN 24 38
FT DOMAIN 39 53
FT DOMAIN 54 60
FT DOMAIN 61 92
FT DOMAIN 93 101
FT DOMAIN 102 111
FT DISULFID 23 92
FT NON_TER 111 111
SQ SEQUENCE 111 AA; 12002 MW; 7A5FCB586C306D29 CRC64;

Query Match 64.1%; Score 373.5; DB 1; Length 111;
Best Local Similarity 67.3%; Pred. No. 3.2e-33;
Matches 76; Conservative 11; Mismatches 23; Indels 3; Gaps 2;

QY 1 DIVMTQSPDSLAVSLGERVTMCKSSQSVLYSSNQKNYLAWYQKPGQSPKLLIYWASTR 60
Db 1 DIVLTQSPASLAVSLGQRATISCKASQSVYDGD--SYMNWYQKPGQPPKLLIYAASNL 58

QY 61 ESGVPDRFSGSGGTFTLTISVQAEDLAVYYCHQYLSS-YTFGGGTKLEIK 112
Db 59 ESGIPARFSGSGGTFTLTINHPVEEDAATYYCQSNEDPFTFGSGTKLEIK 111

RESULT 17
KV30_MOUSE
ID KV30_MOUSE STANDARD; PRT; 111 AA.
AC P01667;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig kappa chain V-III region PC 6308.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE (PC 3741).
RX MEDLINE=79073152; PubMed=103003;
RA Weigert M., Gatmaitan L., Loh E., Schilling J., Hood L.E.;
RT "Rearrangement of genetic information may produce immunoglobulin
diversity.";
RL Nature 276:785-790(1978).
RN [2]
RP SEQUENCE (TEPC 111).
RX MEDLINE=79012520; PubMed=99744;
RA McKean D.J., Bell M., Potter M.;
RT "Mechanisms of antibody diversity: multiple genes encode structurally
related mouse kappa variable regions.";
RL Proc. Natl. Acad. Sci. U.S.A. 75:3913-3917(1978).
CC -I- MISCELLANEOUS: THE PC 3741 AND TEPC 111 SEQUENCES ARE IDENTICAL.
DR PIR; A01934; KVM537.
DR HSSP; P01679; 2FBJ.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IGV; 1.
KW Immunoglobulin V region.
FT DOMAIN 1 23
FT DOMAIN 24 38
FT DOMAIN 39 53
FT DOMAIN 54 60
FT DOMAIN 61 92
FT DOMAIN 93 101
FT DOMAIN 102 111
FT DISULFID 23 92
FT NON_TER 111 111
SQ SEQUENCE 111 AA; 12099 MW; EC46C9D259213BE4 CRC64;

Query Match 63.9%; Score 372.5; DB 1; Length 111;
Best Local Similarity 66.4%; Pred. No. 4.1e-33;
Matches 75; Conservative 15; Mismatches 20; Indels 3; Gaps 2;

QY 1 DIVMTQSPDSLAVSLGERVTMCKSSQSVLYSSNQKNYLAWYQKPGQSPKLLIYWASTR 60
Db 1 DIVLTQSPASLAVSLGQRATISCRASESV--DSYGNFHMHWYQKPGQPPKLLIYRASNL 58

QY 61 ESGVPDRFSGSGGTFTLTISVQAEDLAVYYCHQYLSS-YTFGGGTKLEIK 112
Db 59 ESGIPARFSGSGSRTDFTLTINPVEADDVATYYCQSNEDPYTFGGGTKLEIK 111
```

```
SQ SEQUENCE 111 AA; 12071 MW; 7A4ADE4D6C256D29 CRC64;

Query Match 64.1%; Score 373.5; DB 1; Length 111;
Best Local Similarity 67.3%; Pred. No. 3.2e-33;
Matches 76; Conservative 11; Mismatches 23; Indels 3; Gaps 2;

QY 1 DIVMTQSPDSLAVSLGERVTMCKSSQSVLYSSNQKNYLAWYQKPGQSPKLLIYWASTR 60
Db 1 DIVLTQSPASLAVSLGQRATISCKASQSVYDGD--SYMNWYQKPGQPPKLLIYTASNL 58

QY 61 ESGVPDRFSGSGGTFTLTISVQAEDLAVYYCHQYLSS-YTFGGGTKLEIK 112
Db 59 ESGIPARFSGSGGTFTLTINHPVEEDAATYYCQSNEDPWTFTGSGTKLEIK 111

RESULT 18
KV3H_MOUSE
ID KV3H_MOUSE STANDARD; PRT; 111 AA.
AC P01660;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig kappa chain V-III region PC 3741/TEPC 111.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE (PC 3741).
RX MEDLINE=79073152; PubMed=103003;
RA Weigert M., Gatmaitan L., Loh E., Schilling J., Hood L.E.;
RT "Rearrangement of genetic information may produce immunoglobulin
diversity.";
RL Nature 276:785-790(1978).
RN [2]
RP SEQUENCE (TEPC 111).
RX MEDLINE=79012520; PubMed=99744;
RA McKean D.J., Bell M., Potter M.;
RT "Mechanisms of antibody diversity: multiple genes encode structurally
related mouse kappa variable regions.";
RL Proc. Natl. Acad. Sci. U.S.A. 75:3913-3917(1978).
CC -I- MISCELLANEOUS: THE PC 3741 AND TEPC 111 SEQUENCES ARE IDENTICAL.
DR PIR; A01934; KVM537.
DR HSSP; P01679; 2FBJ.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IGV; 1.
KW Immunoglobulin V region.
FT DOMAIN 1 23
FT DOMAIN 24 38
FT DOMAIN 39 53
FT DOMAIN 54 60
FT DOMAIN 61 92
FT DOMAIN 93 101
FT DOMAIN 102 111
FT DISULFID 23 92
FT NON_TER 111 111
SQ SEQUENCE 111 AA; 12099 MW; EC46C9D259213BE4 CRC64;

Query Match 63.9%; Score 372.5; DB 1; Length 111;
Best Local Similarity 66.4%; Pred. No. 4.1e-33;
Matches 75; Conservative 15; Mismatches 20; Indels 3; Gaps 2;

QY 1 DIVMTQSPDSLAVSLGERVTMCKSSQSVLYSSNQKNYLAWYQKPGQSPKLLIYWASTR 60
Db 1 DIVLTQSPASLAVSLGQRATISCRASESV--DSYGNFHMHWYQKPGQPPKLLIYRASNL 58

QY 61 ESGVPDRFSGSGGTFTLTISVQAEDLAVYYCHQYLSS-YTFGGGTKLEIK 112
Db 59 ESGIPARFSGSGSRTDFTLTINPVEADDVATYYCQSNEDPYTFGGGTKLEIK 111
```



```
RESULT 19
KV3H_HUMAN
ID KV3H_HUMAN STANDARD; PRT; 129 AA.
AC P04207;
DT 20-MAR-1987 (Rel. 04, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig kappa chain V-III region CLL precursor (Rheumatoid factor).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86177570; PubMed=3083417;
RA Jirik F.R., Sorge J., Fong S., Heitzmann J.G., Curd J.G., Chen P.P.,
RA Goldfien R., Carson D.A.;
RT "Cloning and sequence determination of a human rheumatoid factor
RT light-chain gene."
RT Proc. Natl. Acad. Sci. U.S.A. 83:2195-2199(1986).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M12740; AAA58992.1; -.
DR PIR; A01898; K3HUCL.
DR HSSP; P80362; LWTL.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; Ig; 1.
DR Immunoglobulin V region; Signal.
KW SIGNAL
FT CHAIN 1 20 IG KAPPA CHAIN V-III REGION CLL.
FT DOMAIN 21 129 FRAMEWORK-1.
FT DOMAIN 21 43 FRAMEWORK-2.
FT DOMAIN 44 54 COMPLEMENTARITY-DETERMINING-1.
FT DOMAIN 55 69 FRAMEWORK-2.
FT DOMAIN 70 76 COMPLEMENTARITY-DETERMINING-2.
FT DOMAIN 77 108 FRAMEWORK-3.
FT DOMAIN 109 118 COMPLEMENTARITY-DETERMINING-3.
FT DOMAIN 119 129 JK1 SEGMENT.
FT DISULFID 43 108 BY SIMILARITY.
FT NON_TER 129 129
SEQUENCE 129 AA; 14275 MW; 5C13B411BE60CC14 CRC64;

Query Match 63.8%; Score 372; DB 1; Length 129;
Best Local Similarity 64.9%; Pred. No. 5.5e-33;
Matches 74; Conservative 17; Mismatches 15; Indels 8; Gaps 2;

QY 1 DIVMTQSPDSLAVSLGERVTMCKSSQSVLYSSNKNYLAWYQQKPGQSPKLLIYWASTR 60
:||||| :||| ||| :||| :||| :||| :||| :||| :||| :||| :|||
21 EIVMTQSPATLSVSPGERATLSCRASQSV-----SNNLAWYQQKPGQPPRLIYGASTR 74

QY 61 ESGVPDRFSGSGGTFTLTISVQAEDLAVYCHQYLS--SYTFGGGTGLEIK 112
:|:| | | | | | | | | | | | | | | | | | | | | | | | | | | |
75 ATGIPARFSGSGGTFTLTISRLQSEDAFYVYQQYNWPPWTFGGQTRVEIK 128

RESULT 20
KV3N_MOUSE
ID KV3N_MOUSE STANDARD; PRT; 111 AA.
AC P01666;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig kappa chain V-III region PC 7183.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

```
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE.
RX MEDLINE=79073152; PubMed=103003;
RA Weigert M., Gatmaitan L., Loh E., Schilling J., Hood L.E.;
RT "Rearrangement of genetic information may produce immunoglobulin
RT diversity."
RL Nature 276:785-790(1978).
DR PIR; B01937; KVMS83.
DR HSSP; P01679; 2FBJ.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; Ig; 1.
KW Immunoglobulin V region.
FT DOMAIN 1 23 FRAMEWORK-1.
FT DOMAIN 24 38 COMPLEMENTARITY-DETERMINING-1.
FT DOMAIN 39 53 FRAMEWORK-2.
FT DOMAIN 54 60 COMPLEMENTARITY-DETERMINING-2.
FT DOMAIN 61 92 FRAMEWORK-3.
FT DOMAIN 93 101 COMPLEMENTARITY-DETERMINING-3.
FT DOMAIN 102 111 FRAMEWORK-4.
FT DISULFID 23 92 BY SIMILARITY.
FT NON_TER 111 111
SQ SEQUENCE 111 AA; 11952 MW; 2058BB50CE306D31 CRC64;

Query Match 63.0%; Score 367.5; DB 1; Length 111;
Best Local Similarity 66.4%; Pred. No. 1.4e-32;
Matches 75; Conservative 11; Mismatches 24; Indels 3; Gaps 2;

QY 1 DIVMTQSPDSLAVSLGERVTMCKSSQSVLYSSNKNYLAWYQQKPGQSPKLLIYWASTR 60
:||||| | | | | | | | | | | | | | | | | | | | | | | |
1 DIVLTQSPASLAVSLGQRATISCKASQSVVDGDD--SYMNWYQQKPGQPPKLLIYAASNL.58

QY 61 ESGVPDRFSGSGGTFTLTISVQAEDLAVYCHQYLS--YTFGGGTGLEIK 112
|||:| | | | | | | | | | | | | | | | | | | | | | |
59 ESGIPARFSGSGGTFTLTNIHPVEEEDAATYCCQSNEDPLTFTFGAGTKLELK 111

RESULT 21
KV3R_MOUSE
ID KV3R_MOUSE STANDARD; PRT; 111 AA.
AC P01670;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig kappa chain V-III region PC 6684.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE.
RX MEDLINE=79073152; PubMed=103003;
RA Weigert M., Gatmaitan L., Loh E., Schilling J., Hood L.E.;
RT "Rearrangement of genetic information may produce immunoglobulin
RT diversity."
RL Nature 276:785-790(1978).
DR PIR; A01938; KVMS84.
DR HSSP; P80362; LWTL.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; Ig; 1.
KW Immunoglobulin V region.
FT DOMAIN 1 23 FRAMEWORK-1.
FT DOMAIN 24 38 COMPLEMENTARITY-DETERMINING-1.
FT DOMAIN 39 53 FRAMEWORK-2.
FT DOMAIN 54 60 COMPLEMENTARITY-DETERMINING-2.
FT DOMAIN 61 92 FRAMEWORK-3.
FT DOMAIN 93 101 COMPLEMENTARITY-DETERMINING-3.
FT DOMAIN 102 111 FRAMEWORK-4.
```

```
FT DISULFID 23 92 BY SIMILARITY.
FT NON_TER 111 111
SQ SEQUENCE 111 AA; 12039 MW; 1E46988341858526 CRC64;

Query Match
Best Local Similarity 63.0%; Score 367.5; DB 1; Length 111;
Matches 76; Conservative 12; Mismatches 22; Indels 3; Gaps 2;

QY 1 DIVMTQSPDLSAVSLGERVTMCKSSQSVLYSSNQKNYLAHYQKPGQSPKLLIYWASTR 60
Db 1 DIVLTQSPASLAVSLGQRATISCRASKSV--STSGYSYMHWHYQKPGQPCKLLIYLASNL 58

QY 61 ESGVPDRFSGSGGTDTLTITSSVQAEDLAVYYC-HOYLSSYTFGGGKLEIK 112
Db 59 ESGVPARFSGSGGTDTLTINHPVEEDAATYYCQHSRELPRTFGGGKLEIK 111

RESULT 22
KV1W_HUMAN
ID KV1W_HUMAN STANDARD; PRT; 129 AA.
P04431;
13-AUG-1987 (Rel. 05, Created)
13-AUG-1987 (Rel. 05, Last sequence update)
15-JUL-1999 (Rel. 38, Last annotation update)
Ig kappa chain V-I region Walker precursor.
Homo sapiens (Human).
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
[1]
SEQUENCE FROM N.A.
MEDLINE=85014148; PubMed=6091049;
Klobeck H.G., Combriato G., Zachau H.G.;
"Immunoglobulin genes of the kappa light chain type from two human
lymphoid cell lines are closely related.";
Nucleic Acids Res. 12:6995-7006(1984).
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; X00965; CAA25477.1; ALT_TERM.
PIR; A01883; KLHUK.
HSSP; P01607; IREI.
InterPro; IPR003006; Ig_MHC.
InterPro; IPR003596; Ig_v.
Pfam; PF00047; ig; 1.
SMART; SM00406; IGV; 1.
Immunoglobulin v region; Signal.
KW SIGNAL 1 22
CHAIN 23 129 IG KAPPA CHAIN V-I REGION WALKER.
DOMAIN 23 45 FRAMEWORK-1.
DOMAIN 46 56 COMPLEMENTARITY-DETERMINING-1.
DOMAIN 57 71 FRAMEWORK-2.
DOMAIN 72 78 COMPLEMENTARITY-DETERMINING-2.
DOMAIN 79 110 FRAMEWORK-3.
DOMAIN 111 119 COMPLEMENTARITY-DETERMINING-3.
DOMAIN 120 129 FRAMEWORK-4.
DISULFID 45 110 BY SIMILARITY.
NON_TER 129 129
SEQUENCE 129 AA; 14069 MW; F941FA07D4AFC2F9 CRC64;

Query Match
Best Local Similarity 63.0%; Score 367.5; DB 1; Length 129;
Matches 74; Conservative 14; Mismatches 18; Indels 7; Gaps 2;

QY 1 DIVMTQSPDLSAVSLGERVTMCKSSQSVLYSSNQKNYLAHYQKPGQSPKLLIYWASTR 60
Db 23 DIQMTQSPSSLSASVGRVITITCRASQSI-----SNYLNWYQKPGKAPKLLIYAASSL 76
```

```
QY 61 ESGVPDRFSGSGGTDTLTITSSVQAEDLAVYYCHQ-YLSSYTFGGGKLEIK 112
Db 77 QSGVTSRFSGSGGTDTLTITSSLPQEDSATYYCQSYSTLITFGQGRLEIK 129

RESULT 23
KV3A_HUMAN
ID KV3A_HUMAN STANDARD; PRT; 108 AA.
P01619;
21-JUL-1986 (Rel. 01, Created)
21-JUL-1986 (Rel. 01, Last sequence update)
15-JUL-1999 (Rel. 38, Last annotation update)
Ig kappa chain V-III region B6.
Homo sapiens (Human).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
[1]
SEQUENCE.
RA Milstein C.;
"The basic sequences of immunoglobulin kappa chains: sequence studies
of Bence Jones proteins Rad, Fr4 and B6.";
FEBS Lett. 2:301-304(1969).
CC -!- MISCELLANEOUS: THIS IS A BENCE-JONES PROTEIN.
DR PIR; A01891; K3HUB6.
DR HSSP; P80362; 1WTL.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; ig; 1.
DR SMART; SM00406; IGV; 1.
KW Immunoglobulin V region; Bence-Jones protein.
FT DISULFID 23 89 BY SIMILARITY.
FT NON_TER 108 108
SQ SEQUENCE 108 AA; 11635 MW; 8BC14FF07A419E3D CRC64;

Query Match
Best Local Similarity 63.0%; Score 367; DB 1; Length 108;
Matches 69; Conservative 22; Mismatches 16; Indels 6; Gaps 2;

QY 1 DIVMTQSPDLSAVSLGERVTMCKSSQSVLYSSNQKNYLAHYQKPGQSPKLLIYWASTR 60
Db 1 ZIVLTZSPGTLSPGZRAALSCRASQSL-----SGNYLAHYQKPGQAPRLLMYGVSSR 55

QY 61 ESGVPDRFSGSGGTDTLTITSSVQAEDLAVYYCHQYLSS-YTFGGGKLEIK 112
Db 56 ATGIPDRFSGSGGTDTLTISRLZPDEFVAVYCCQYGGSPFTFGGSKLEIK 108

RESULT 24
KV3F_HUMAN
ID KV3F_HUMAN STANDARD; PRT; 109 AA.
P01624;
21-JUL-1986 (Rel. 01, Created)
21-JUL-1986 (Rel. 01, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
Ig kappa chain V-III region POM.
Homo sapiens (Human).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
[1]
SEQUENCE.
RX MEDLINE=76276460; PubMed=60899;
RA Klapper D.G., Capra J.D.;
"The amino acid sequence of the variable regions of the light chains
from two idiotypically cross reactive IgM anti-gamma globulins.";
Ann. Immunol. (Paris) 127C:261-271(1976).
CC -!- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM AN IGM WITH ANTI-GAMMA
GLOBULIN ACTIVITY.
DR PIR; A01897; K3HUPM.
DR HSSP; P80362; 1WTL.
DR InterPro; IPR003006; Ig_MHC.
```


